QUALITY ASSURANCE PROJECT PLAN

for the

Maritime Environmental Resource Center

Version 1.0

31 January 2012

University of Maryland Center for Environmental Science
Chesapeake Biological Laboratory
One Williams Street
Solomons, MD 20688

A. PROJECT MANAGEMENT

A.1. Title and Approvals

Title: Quality Assurance Project Plan for the Maritime Environmental Resource Center, Version 1.0

Signed by Carolynn Junemann

MARAD PROJECT OFFICER **Date**

Signed by Frank Hamons

MARYLAND PORT ADMINISTRATION **Date**

Signed by Mario Tamburri

MERC DIRECTOR **Date**

Signed by Earle Buckley

MERC QUALITY ASSURANCE MANAGER **Date**

Maritime Environmental Resource Center
Chesapeake Biological Laboratory
One Williams Street
Solomons, MD 20688

A.2. Table of Contents

LIST OF FIGURES

LIST OF TABLES

APPENDICES

A.3. Distribution List

This list includes the names and contact information of those who receive copies of the approved MERC Program QAPP and subsequent revisions.

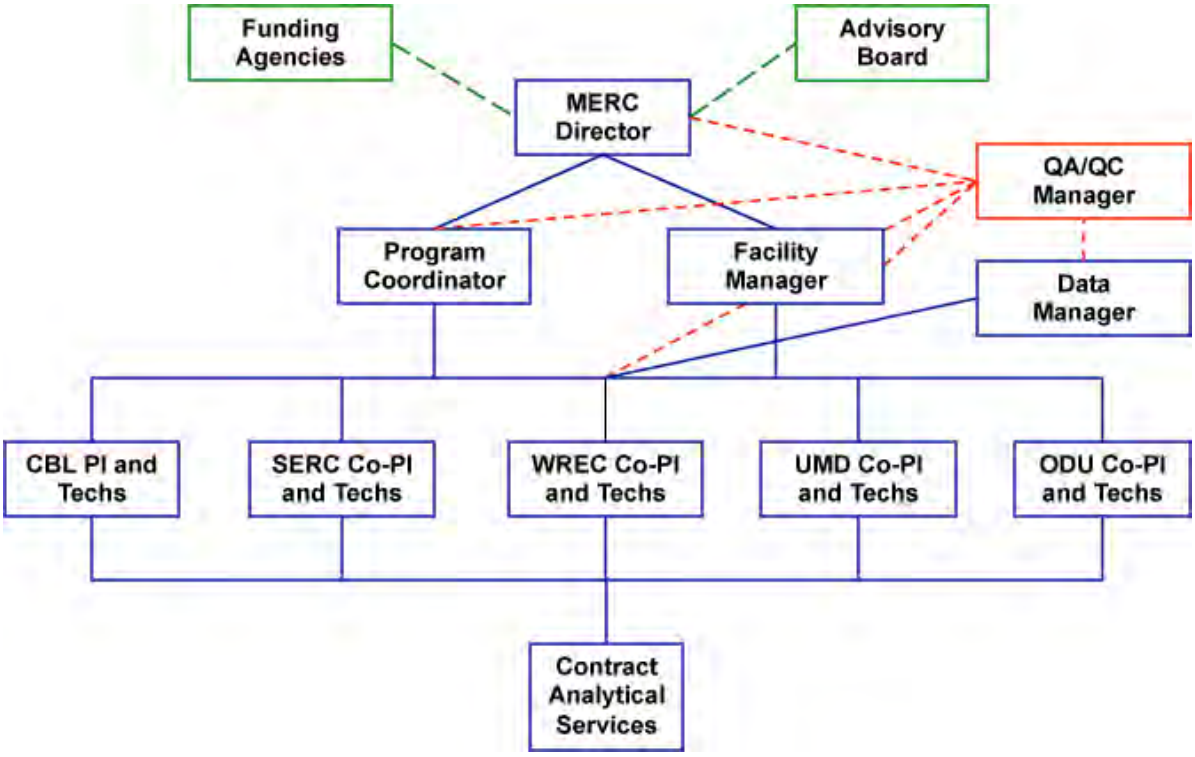
Name	Position	Organization	Contact Information
Dr. Mario Tamburri	MERC Director and Principal Investigator	UMCES/CBL	410-326-7440 tamburri@umces.edu
Ms. Janet Barnes	MERC Program Coordinator	UMCES/CBL	410-326-7259 barnes@umces.edu
Mr. George Smith	MERC Facility Manager	SERC	443-482-2411 smithgeo@si.edu
Ms. Katherine Ziombra	MERC Data Manager	UMCES/CBL	410-385-6311 davis@umces.edu
Dr. Earle Buckley	MERC Quality Assurance Manager	Buckley Environmental, LLC	843-991-2751 earlebuckley@comcast.net
Dr. Gregory Ruiz	MERC Partner Institution lead and Co-PI	SERC	443-482-2227 ruizg@si.edu
Dr. Daniel Fisher	MERC Partner Institution lead and Co-PI	WREC	410-827-8056 dfisher2@umd.edu
Dr. Anwar Huq	MERC Partner Institution lead and Co-PI	UMD	301-405-7428 huqanwar@gmail.com
Dr. Fred Dobbs	MERC Partner Institution lead and Co-PI	ODU	757-683-5329 fdobbs@odu.edu.
Dr. Carolyn Junemann	Environmental Protection Specialist	MARAD	202-366-1920 carolyn.junemann@dot.gov
Mr. Frank Hamons	Deputy Director Harbor Development	Maryland Port Administration	410-385-4445 fhamons@marylandports.com

A.4. MERC Organization

The University of Maryland Center for Environmental Science's Chesapeake Biological Laboratory (CBL) is the lead for MERC. Other MERC testing and research partners include the Smithsonian Environmental Research Center (SERC), University of Maryland College Park (UMD), University of Maryland Wye Research and Education Center (WREC), and Old Dominion University (ODU). MERC sponsors include the Maryland Port Administration (MPA), US Maritime Administration (MARAD), and National Oceanic and Atmospheric Administration (NOAA).

The organization chart in Figure 1 identifies the responsibilities of the organizations and individuals associated with these ballast water treatment systems (BWTS) tests. Roles and responsibilities for the BWTS testing described in this document are defined further below.

Figure 1. MERC Organization



A.4.1. Director and Principal Investigator

Dr. Mario Tamburri of CBL is the MERC Director and Principal Investigator and has the final authority on decisions related to MERC BWTS tests. Dr. Tamburri will:

- review the draft and approve the final QAPP and Test Plans;
- review the draft and approve final test reports;
- ensure that necessary MERC resources, including staff and facilities, are committed to the BWTS tests;
- ensure that the technical, schedule, and cost goals established for the tests are met;
- ensure that confidentiality of sensitive vendor information is maintained;
- ensure that testing staff respond to QAPP and Test Plans deviations and any issues raised in assessment reports, audits, or from test staff observations, and that any necessary corrective actions have been implemented;
- facilitate a stop work order if the MERC QA Manager discovers adverse findings that will compromise data quality or test results.

A.4.2. Program Coordinator

Janet Barnes at CBL/UMCES is the MERC Program Coordinator (PC) and coordinator for BWTS tests. In this role, Ms. Barnes will:

- prepare the draft and final Test Plans;
- coordinate distribution of final Test Plans;

- coordinate a kick-off meeting prior to the start of the tests to review the critical logistical, technical, and administrative aspects of the tests and confirm responsibilities;
- serve as the primary point of contact for manufacturer representatives;
- coordinate the MERC team to conduct the BWTS performance tests in accordance with the MERC Quality Management Plan (QMP) and this Quality Assurance Project Plan (QAPP);
- maintain real-time communication with the MERC Director and QA Manager on any potential or actual deviations from the QAPP or a specific Test Plan;
- respond to any issues raised in assessment reports and audits, including instituting corrective action as necessary;
- prepare the draft test report and revise in response to reviewers' comments;
- coordinate distribution of the final test report.

A.4.3. Facility Manager

George Smith, a SERC Biological Research Technician, is the MERC Facillity Manager. And in this role, is responsible for the

- operations, maintenance and/or modification to the Mobile Test Platform
- works closely with the Director and senior scientists to assure effective sample collection and handling at the Mobile Test Platform.
- works with the PC on developing SOPs as they relate to operation of the facility and is responsible for ensuring worker health and safety at the site. The Facility Manager is responsible for all final decisions made during testing and supports data compilation, analysis and reporting in consultation with the Director and Program Coordinator, the

.

A.4.4. Test Team Members

Researchers from University of Maryland Center for Environmental Science's Chesapeake Biological Laboratory (CBL), Smithsonian Environmental Research Center (SERC), University of Maryland College Park (UMD), University of Maryland Wye Research and Education Center (WREC) and Old Dominion University (ODU) are members of the MERC testing team. These researchers' responsibilities include:

- assist in developing Test Plans;
- perform sample collections and analyses in their specific area of expertise as described in the QAPP and Test Plans;
- provide all test data to the Data Manager electronically;
- provide copies of all field and lab sheets to Data Manager;
- provide the QA Manager with all appropriate QA documentation of their respective test facilities, including equipment and procedures;
- respond to any issues raised in assessment reports and audits, including instituting corrective action as necessary;
- prepare sections of the draft test report relevant to their specific area of responsibility.

Dr. Gregory Ruiz, Senior Scientist at SERC and head the Marine Invasion Research Laboratory, is responsible for working with MERC Director on SOPs for sample collections, live counts of organisms > 50 μm and live counts for organisms 10 – 50 μm, and oversees SERC Biological Research Technicians involved in these activities, George Smith, Timothy Mullady and Darrick Sparks.

Dr. Anwar Huq, Maryland Pathogen Research Institute (MPRI), University of Maryland, is MERC's Senior Microbialist. Dr. Huq is responsible for developing the microbial-related SOPs, review test

results, and supervises Dr. Elisa Taviani and Graduate Research Assistants at MPRI to assure appropriate microbial sample collection and handling, and analysis of microbial samples according to relevant SOPs.

Dr. Daniel Fisher, Director of Maryland Department of the Environment's Bioassay Laboratory located at WREC, is responsible for developing SOPs for toxicity testing and residual byproduct chemical analyses, review test results, and oversees WREC staff involved in these activities, including Dr. Lance Yonkos, Assistant Research Scientist, and Gregory Ziegler, Faculty Research Associate.

Dr. Fred Dobbs, Professor and Graduate Program Director,Department of Ocean, Earth and Atmospheric Sciences at Old Dominion University, is responsible for scientific review of all SOPs and test results, and for logistic support for MERC testing in Norfolk, VA.

Katherine Davis Ziombra, at CBL/UMCE,S is the Data Manager and is responsible for the compilation, review, management and storage of all data collected during MERC testing. The Data Manager will coordinate with the Quality Assurance Manager on revisions to the data.

A.4.5. Quality Assurance Manager

Dr. Earle Buckley is the Quality Assurance (QA) Manager for MERC and provides independent oversight of the MERC quality system. For MERC BWTS testing, Dr. Buckley will:

- review the draft and final QAPP and Test Plans;
- attend the BWTS test kick-off meeting and lead the discussion of the QA elements of the testing;
- prior to the start of testing, verify the presence of applicable training records, including any vendor training on test equipment;
- prepare audit checklists;
- conduct a technical systems audit at least once near the beginning of each BWTS test;
- conduct audits to verify data quality;
- prepare and distribute an audit report for each audit;
- verify that audit responses for each audit finding and observation are appropriate and that corrective action has been implemented effectively;
- maintain real-time communication with the PC on QA activities, audit results, and concerns,
- communicate to the PC and/or technical staff the need for immediate corrective action if an audit identifies QAPP and/or Test Plan deviations or practices that threaten data quality;
- recommend a stop work order if audits indicate that data quality or safety is being compromised;
- work with the PC and MERC Director to resolve data quality concerns and disputes;
- provide a summary of the QA/quality control (QC) activities and results for the final reports;
- review the draft and final test reports;
- review and approve QAPP and Test Plan amendments and deviations.

A.4.6. Manufacturer

Manufacturer representatives will:

- review the draft Test Plan and provide comments and recommendations;
- approve the final Test Plan;
- interface with the MERC PC to make all arrangements for the test;
- sign a MERC manufacturer agreement to participate in the test;
- provide operational treatment systems for the agreed upon test site(s) for the duration of the test;

- commit a trained technical representative to install and operate, maintain, and repair the treatment system throughout the test or train an operator to perform these tasks and sign a consent form indicating training occurred;
- inspect the installation and operation of the system prior to the initiation of the testing;
- review their respective draft test reports.

A.5. Background

A.5.1. Problem Definition

Invasions by non-native aquatic species are increasingly common worldwide, often causing ecological and economic damage, and it is widely accepted that ballast water is one of the most important vectors for transporting and introducing non-native species to new biogeographic regions. Consequently, the International Maritime Organization (IMO), through the 2004 International Convention for the Control and Management of Ships' Ballast Water and Sediment, and more recently the proposed US Coast Guard (USCG) and Environmental Protection Agency (US EPA) Vessel General Permit (VGP) have all put forward similar ballast water discharge standards that limit concentrations of living organisms in different size or taxonomic categories that can be released with ballast water. Current proposed ballast water discharge standards include:

- Less than 10 viable organisms per one m^3 greater than or equal to 50 µm in minimum dimension
- Less than 10 viable organisms per ml less than 50 µm in minimum dimension and greater than or equal to 10 µm in minimum dimension
- Less than the following concentrations of indicator microbes, as a human health standard: a) Toxicogenic *Vibrio cholerae* (serotypes O1 and O139) with less than 1 colony forming unit (cfu) per 100 ml; b) *Escherichia coli* less than 250 cfu per 100 ml; and c) intestinal *Enterococci* less than 100 cfu per 100 ml.

To address the IMO and US discharge standards, technology developers and manufacturers around the world have designed and built a variety of onboard ballast water treatment systems (BWTSs) to achieve the prescribed discharge limits. A BWTS is a "Prefabricated, commercial-ready, treatment systems designed to remove, kill or inactivate (prior to discharge) organisms in ballast water. This includes all components, in an integrated fashion, required for shipboard operation." (EPA, 2010).

Prior to any approval or certification of a BWTS, all systems must go through extensive phased development and testing from the laboratory to full-scale shipboard verification. This phased approach not only addresses engineering challenges of scaling up, but also develops a comprehensive understanding of the system's mode of action (i.e., how the treatment kills or removed organisms) and the dose-response of various organisms to a range of treatment conditions. This knowledge and testing can be used to help identify indirect measures of system efficacy and compliance. Three documents have been produce to provide guidance and standardization of BWTS testing:

- International Maritime Organization (2005) Resolution MEPC.125(53) Guidelines for Approval of Ballast Water Management Systems (G8);
- International Maritime Organization (2008) Resolution MEPC.125(57) Revised Procedure for Approval of Ballast Water Management Systems that Make Use of Active Substances (G9); and
- ETV Generic Protocols for the Verification of Ballast Water Treatment Technologies, (2010) EPA/600/R-10/146 (EPA).

Administrations and Classification Societies utilize the results for these series of land-based and shipboard BWTS tests, combined with other relevant information, for a final decision on Type Approval Certification of individual BWTS.

A.5.2. MERC Background

MERC was created by the University of Maryland Center for Environmental Science and Maryland Port Administration, with additional support from the US Maritime Administration, and National Oceanic and Atmospheric Administration, to provide test facilities, expertise, information, and decision tools to address key environmental issues facing the international maritime industry. The primary focus is to evaluate the mechanical and biological efficacy, costs, and logistical aspects of ballast water treatment systems and to assess the economic impacts of ballast water regulations and management approaches.

The goal of MERC BWTS testing is to conduct independent, scientifically-sound, quality-assured evaluations of treatment approaches and systems with regard to factors such as biological treatment efficacy, predictability/reliability, environmental acceptability, and safety. MERC conducts R&D and certification testing of treatment systems at three levels: lab bench proof-of-concept, land-based prototype, and shipboard validation/verification. All MERC testing protocols are based on the IMO G8 and G9 Guidelines and the ETV Protocol, and employ scientifically validated or accepted approaches methods.

While the initial and primary focus of MERC is on ballast water treatment systems, the Center has the expertise, facilities, academic independence, and scientific integrity that will allow for testing and assessment of additional technologies and innovations related to Green Shipping, including hull fouling invasive species, port and vessel air emissions and alternative fuels, and gray and oily water treatments.

A.5.3. MERC BWTS Test Objectives

MERC's four main objectives are:

- provide technology developers/vendors with facilities and expertise for pilot-scale and shipboard evaluations of treatment systems;
- provide regulatory agencies and classification societies with standardized, rigorous, and independent data on treatment system performance;
- provide ship builders and shipping lines with information and decision tools to select the most appropriate ballast water treatment options; and
- remove as much uncertainty as possible from emerging markets for treatment systems in order to accelerate the adoption of innovative technologies.

A.6. Test Descriptions and Schedule

MERC test activities may have several goals including:

- Pre-certification testing, i.e., operational and biological performance (including residual toxicity) status-testing given scale-up and a range of challenge conditions; and
- Certification/verification testing, i.e., formal assessment of performance against IMO, USCG and other discharge standards.

The fundamental approach of MERC testing is to conduct independent, scientifically-sound, rigorous, and quality assured evaluations of ballast water treatment system performance under controlled experimental conditions. In addition, MERC tests are directly relevant to regulatory processes including the IMO Convention, state law, and federal requirements under development in the United States. To that end, MERC protocols, challenge conditions and testing infrastructure (e.g. flow rate, retention tank size, sample size, sample collection and analysis equipment and data logging) are based on the essential

features of the IMO G8 guidelines for testing, and the ETV protocols. MERC testing also can be adapted to address other possible benchmarks such as stricter performance standards or non-regulatory end-points.

A.6.1. Test Description

Systems that will be tested under MERC will be capable of treating the entire discharge or ballast water volume for biological organisms, either through a one-step treatment process or through multi-step treatment processes, and will be capable of treating a wide range of source water typical of ballast uplifted from fresh, coastal, estuarine and/or marine origins. These technologies may be mechanical, chemical, physical or biological in nature or a combination of any of the technologies. Treatment systems or components of systems that provide only partial treatment of the discharge can be evaluated as separate demonstration exercises but are excluded from full certification testing. The factors that are verified during BWTS testing include: biological treatment performance, operation and maintenance, predictability/reliability, cost factors, environmental acceptability, and safety. In performing a BWTS test MERC follows the technical and QA procedures specified in this QAPP and complies with the data quality requirements in the MERC QMP. Table 1 compares MERC test protocols with those of IMO G8 Guidelines and ETV Protocols.

Table 1. Comparison of Key Test Parameters Proposed for MERC Tests with G8 and ETV Test Parameters.

Parameter	Sub-category	G8	ETV	MERC
Organisms To Be Evaluated	Zooplankton, live organisms ≥ 50 μm in size	Naturally occurring, or cultured organisms may be added to the test water.	Ambient assemblage supplemented by the addition of standard test organisms.	Naturally occurring Chesapeake Bay assemblage. Native culture organisms can be added if required.
	Protists, live organisms 10 - 50 μm in size	Naturally occurring, or cultured species that may be added to the test water.	Ambient assemblage supplemented by the addition of standard test organisms.	Naturally occurring Chesapeake Bay assemblage. Culture organisms can be added if required.
	Bacteria	Naturally occurring, or cultured species that may be added to the test water.	Ambient assemblage supplemented by the addition of standard test organisms.	Naturally occurring Chesapeake Bay assemblage.
Intake Organism Diversity & Density	Zooplankton, live organisms ≥ 50 μm in size	Organisms ≥50 μm in minimum dimension should be present in a total density of preferably 10^6 individuals but not less than 10^5 individuals per m3, and should consist of at least 5 species from at least 3 different phyla/divisions.	Total concentration = minimum of 1 x 10^5 organisms/m^3.	Organisms ≥ 50 μm in minimum dimension are typically present in a total density above 10^5 live individuals per m^3, and consist of at least 5 species from at least 3 different phyla/divisions.
	Protists, live organisms 10 - 50 μm in size	Organisms ≥10 μm and less than 50 μm in minimum dimension should be present in a total density of preferably 10^4 individuals but not less than 10^3 individuals per ml, and should consist of at least 5 species from at least 3 different phyla/divisions.	Organisms in the ≥10 μm and <50 μm size class must be present in minimum concentrations of 10^3 organisms/ml with at least 5 species across 3 phyla.	Entities ≥10 μm and less than 50 μm in minimum visible dimension are typically present in a total density above 10^3 cells per ml, and consist of at least 5 species from at least 3 different phyla.
	Bacteria	Heterotrophic bacteria should be present in a density of at	Organisms in the < 10 μm size class must be present	Heterotrophic bacteria are typically present in a

Parameter	Sub-category	G8	ETV	MERC
		least 10^4 living bacteria per ml.	in minimum concentrations of 10^3/ml as culturable aerobic heterotrophic bacteria.	density of at least 10^3/ml as culturable aerobic heterotrophic bacteria.
Water Quality of Intake/ Source Water	N/A	• Dissolved Organic Carbon (DOC): >5 mg/l; • Particulate Organic Carbon (POC): >5 mg/l; • Total Suspended Solids (TSS): >50 mg/l.	• Dissolved Organic Matter (DOM): min. 6 mg/l as DOC; • Particulate Organic Matter (POM): min. 4 mg/l as POC; • Mineral Matter (MM): min. 20 mg/l; • Total Suspended Solids (TSS): = POM + MM: min. 24 mg/l;	Dependent season and location, typical ambient values include: • Dissolved Organic Carbon (DOC): 2-5 mg/l; • Particulate Organic Carbon (POC): 0.5 - 2 mg/l; • Total Suspended Solids (TSS): 10 - 25 mg/l. Typically augmented to increase levels.
Salinity of Intake/Source Water		• Freshwater <3 PSU; • 10 PSU difference to brackish and marine	• Fresh <1 PSU; • Brackish 10 - 20 PSU • Marine 28 – 36 PSU	Dependent season and location, typical ambient values include: • Anacostia River Washington DC <1 PSU • Baltimore Harbor 5 – 12 PSU • Norfolk Harbor 19 – 25 PSU
Sample Volume	Zooplankton, live organisms ≥ 50 µm in size	At least 20 l of intake water and 1 m^3 of treated water.	Minimum of 3 m^3 concentrated to 1000 ml per sample.	Between 3 and 10 m^3, concentrated to approximately 1000 ml per sample.
	Protists, live organisms 10 - 50 µm in size	At least 1 l of intake water and 10 l of treated water.	Minimum of 3 m^3 concentrated to 1000 ml per sample.	At least 3 l per time integrated sample.
	Bacteria	At least 500 ml of intake water and 500 ml of treated water.	1000 ml per sample.	At least 3 l per time integrated sample
Number of Intake Samples	Zooplankton, live organisms ≥ 50 µm in size	Minimum of 3 samples collected from the treatment track and 3 samples collected from the control track.	1 sample immediately prior to water entry to the control tank and 1 sample immediately before entry to the in-line BWTS, or (if control and challenge water are shown to be representative) one sample before the splitter.	1 continuous time-integrated sample collected from the control and post treatment lines at uptake. Split for analysis.
	Protists, live organisms 10 - 50 µm in size	Minimum of 3 samples collected from the treatment track and 3 samples collected from the control track.	1 sample immediately prior to water entry to the control tank and 1 sample immediately before entry to the in-line BWTS, or (if control and challenge water are shown to be representative) one sample before the splitter.	1 continuous time-integrated 90 l sample collected from the control and post treatment lines at uptake. Representative samples analyzed.
	Bacteria	Minimum of 3 samples collected from the treatment track and 3 samples collected from the control track.	1 sample immediately prior to water entry to the control tank and 1 sample immediately before entry to the in-line BWTS, or (if control and challenge water	1 continuous time-integrated sample collected from the control and post treatment lines at uptake. Representative samples analyzed.

Parameter	Sub-category	G8	ETV	MERC
			are shown to be representative) one sample before the splitter.	
Number of Discharge Samples	Zooplankton, live organisms ≥ 50 µm in size	Minimum of 3 samples collected from the treatment track and 3 samples collected from the control track.	1 sample from the discharge of the control tank, and 1 sample from the discharge (following any treatments) of the treated water.	1 continuous time-integrated sample collected from the control and treatment lines upon discharge. Control split for analysis. Whole treatment sample examined.
	Protists, live organisms 10 - 50 µm in size	Minimum of 3 samples collected from the treatment track and 3 samples collected from the control track.	1 sample from the discharge of the control tank, and 1 sample from the discharge (following any treatments) of the treated water.	1 continuous time-integrated sample collected from the control and treatment lines upon discharge. Representative samples analyzed.
	Bacteria	Minimum of 3 samples collected from the treatment track and 3 samples collected from the control track.	1 sample from the discharge of the control tank, and 1 sample from the discharge (following any treatments) of the treated water.	1 continuous time-integrated sample collected from the control and treatment lines upon discharge. Representative samples analyzed.
Analytic Endpoints: Discharge Density	Zooplankton, live organisms ≥ 50 µm in size	Less than 10 viable organisms per m3 greater than or equal to 50 µm in minimum dimension for treated water; more than 100 viable organisms per m3 greater than or equal to 50 µm in minimum dimension for control water.	Treatment efficacy will be determined by the measurement of living ambient organism concentrations in the treatment discharge. Minimum concentration in control tank discharge is 100 live organisms/m3.	Dependent on test plan. May include control vs treatment, intake vs discharge, treatment discharge vs regulatory standard (i.e., IMO and/or ETV).
	Protists, live organisms 10 - 50 µm in size	Less than 10 viable organisms per mL less than 50 µm in minimum dimension and greater than or equal to 10 µm in minimum dimension for treated water; more than 100 viable organisms per mL less than 50 µm in minimum dimension and greater than or equal to 10 µm in minimum dimension for control water.	Treatment efficacy will be determined by the measurement of living ambient organism concentrations in the treatment discharge. Minimum concentration in control tank discharge is 100 live organisms/mL.	Dependent on test plan. May include control vs treatment, intake vs discharge, treatment discharge vs regulatory standard (i.e., IMO and/or ETV).
	Bacteria	Less than 1 colony forming unit (cfu) per 100 mL or less than 1 cfu / 1 g (wet weight) zooplankton of Toxicogenic *Vibrio cholerae* (O1 and O139), less than 250 cfu / 100 mL of E. coli, and less than 100 cfu / 100 mL of intestinal *Enterococci* for treated water; more than 10 cfu / 100 mL or more than 10 cfu / 1 g (wet weight) zooplankton of Toxicogenic Vibrio cholerae (O1 and O139), more than 2500 cfu / 100 mL of *E. coli*, and more than 1000 cfu / 100 mL of intestinal *Enterococci* for control water.	Treatment efficacy will be determined by the measurement of living ambient organism concentrations in the treatment discharge. Minimum concentration in control tank discharge is 5 x 102/mL.	Dependent on test plan. May include control vs treatment, intake vs discharge, treatment discharge vs regulatory standard (i.e., IMO and/or ETV).

Parameter	Sub-category	G8	ETV	MERC
Water Quality Measurements	N/A	pH, temperature, salinity, dissolved oxygen, TSS, DOC, POC and turbidity (NTU) should be measured at the same time that the samples are collected.	Temperature, salinity, TSS, POM, DOM, mineral matter, dissolved oxygen, pH, chlorophyll a.	Dependent on test plan. May include salinity, DOC, POC; Mineral Matter; TSS, dissolved oxygen; temperature, pH, total chlorophyll and others as required.
Toxicity	N/A	Separate samples should be collected for toxicity testing of treated water, from the discharge, for systems that make use of Active Substances and also for those, which could reasonably be expected to result in changes to the chemical composition of the treated water such that adverse impacts to receiving waters might occur upon discharge. Tests should conducted in accordance with paragraphs 5.2.3 to 5.2.7 of the Procedure for Approval of Ballast Water Management Systems That Make Use of Active Substances (resolution MEPC.126(53)) as amended .	Toxicity tests will be conducted for treatments involving biocides. Tests will be selected from a short list of U.S. EPA standard tests.	Dependent on test plan, e.g., Whole effluent toxicity (WET) tests and residual byproduct chemical analyses using treatment discharge water for systems involving active substances.
Biological Sample Analysis	N/A	Samples should be analyzed as soon as possible after sampling, and analyzed live within 6 hour or treated in such a way as to ensure that proper analysis can be performed. Widely accepted standard methods for the collection, handling, storage, and analysis of samples should be used.	Zooplankton enumeration: Concentrate using 35 µm mesh plankton nets; no preservation; sub-sample into well plate (20 1mL wells observed); observe with dissecting microscope and probe organisms to determine live/dead status; fix with Lugol's for total counts. Phytoplankton enumeration: No preservation; stain with Fluorescein Diacetate (FDA) and CMFDA; load into a Sedgewick Rafter Counting Chamber and examine under epifluorescence using a FITC narrow pass filter cube. Bacteria: Plate on appropriate media; use a DNA colony blot hybridization for V. cholerae.	Direct counts (number live) for treatment discharge samples, indirect counts (number of dead and total) for intake and control discharge samples for organisms in the >50 µm size class; direct counts (number of live) using FDA+CMFDA vital stain for organisms in the 10-50 mm size class; enumeration (using appropriate media) of total viable heterotrophic bacteria, *E. coli*, and Enterococci and preparation of colony blots for the detection of toxigenic *Vibrio cholerae*.
Flow Rate	N/A	At least 200 m^3/hr.	At least 200 m^3/hr.	Up to 350 m^3/hr and no lower than 100 m^3/hr.

Parameter	Sub-category	G8	ETV	MERC
Number and Capacity of Retention Tanks	N/A	At least 1 control and 1 treatment tank with a minimum capacity of 200 m^3 each.	At least 1 control and 1 treatment tank with a minimum capacity of 200 m^3 each.	1 control and 1 treatment tanks each with a capacity of 310 m^3.
Control/ Treatment Cycle Sequence	N/A	Control and treatment cycles may be run simultaneously or sequentially.	Control and treatment cycles may be run simultaneously or sequentially.	Control and treatment cycles run simultaneously on uptake and sequentially on discharge
Retention Time	N/A	At least 5 days.	Minimum of one day.	1 to 5 days, dependent on test plan.
Number of Trials	N/A	At least 5 successes.	Minimum of three per salinity regime.	Minimum of 3 trials dependent on test plan.
QA/QC	N/A	Quality Management Plan (QMP) addressing the quality control management structure and policies of the testing body, including subcontractors and outside laboratories; Quality Assurance Project Plan (QAPP) addressing the specifics of the ballast treatment technology to be tested, the test facility, and other conditions affecting the actual design and implementation of the required experiments.	A Test/ Plan (TQAP), also called a QAPP, is to be compiled by the Testing Organization, with input from the vendor. The TQAP will describe the procedures for conducting a test or study according to the verification protocol requirements for the application of a ballast water treatment system at a particular site. At a minimum, the TQAP shall detail test objectives, specific test procedures (including sample and data collection, sample handling, analysis and preservation), and quality control and assurance requirements (including measures of precision, accuracy, comparability, and representativeness).	Quality Management Plan (QMP) addressing the quality control management structure and policies of MERC; Quality Assurance Project Plan (QAPP) addressing the specifics of the MERC's ballast treatment tests, its facilities, and other conditions affecting the actual design and implementation of the required experiments. A Test Plan, to be compiled with input from the vendor.

A.6.2. Test Implementation Schedule

Table 2 shows a general schedule of testing and data analysis/reporting activities to be conducted in MERC BWTS testing. Actual dates are specified in each Test Plan.

Table 2. General MERC BWTS Test schedule

Approximate Months after Start Date	Testing Activities	Data Analysis and Reporting
0 to 3	• Test plan development and approval • Conduct pre-test checks and dry runs	• Not Applicable
3	• Coordinate for technologies and testing supplies to be delivered to test sites • Install necessary equipment and technology • Technology training by manufacturer or coordination with manufacturer representative	• Prepare report template
3 to 5	• Perform Performance Evaluation Audit (PEA) • Complete PEA report • Conduct performance testing • Perform TSA • Perform initial ADQ (1st batch) • Complete performance testing • Perform second ADQ (20% of all data)	• Compile PEA results • Compile data • Review and summarize data • Perform data analysis • Begin draft reports
5	• Prepare draft test report(s) • Perform third ADQ of report(s)	• Complete draft test report(s) • Complete internal review of draft report(s)
6 to 7	• Coordinate reviews of draft report(s)	• Complete peer review and vendor review of draft report(s)
8	• Prepare final test report(s)	• Revise draft test report(s) • Distribute final reports(s) to MPA, MARAD and manufacturers

A.6.3. Test Site Descriptions

MERC offers testing on a Mobile Test Platform, that will allow ballast water treatment systems to be evaluated in Baltimore Harbor, MD (salinity 5 - 12 psu), Norfolk, VA (salinity 19 - 25 psu) and/or Washington, DC (Anacostia River, 0 psu) with one system installation (Figure 2). The MERC facility (Figure 3) includes two 310 m^3 tanks, two 350 m^3/hr pumps and all other associated infrastructure to complete G8/G9 and ETV testing. With respect to challenge conditions, uptake water from these test sites commonly meets many of the IMO G8 and ETV requirements for intake organism densities and physical/chemical conditions during the testing season. For example, Table 3 shows the historical range of biological, physical and chemical for Baltimore Harbor in comparison with G8 guidelines and ETV protocols testing conditions.

Figure 2. Typical salinity ranges for the Chesapeake Bay.

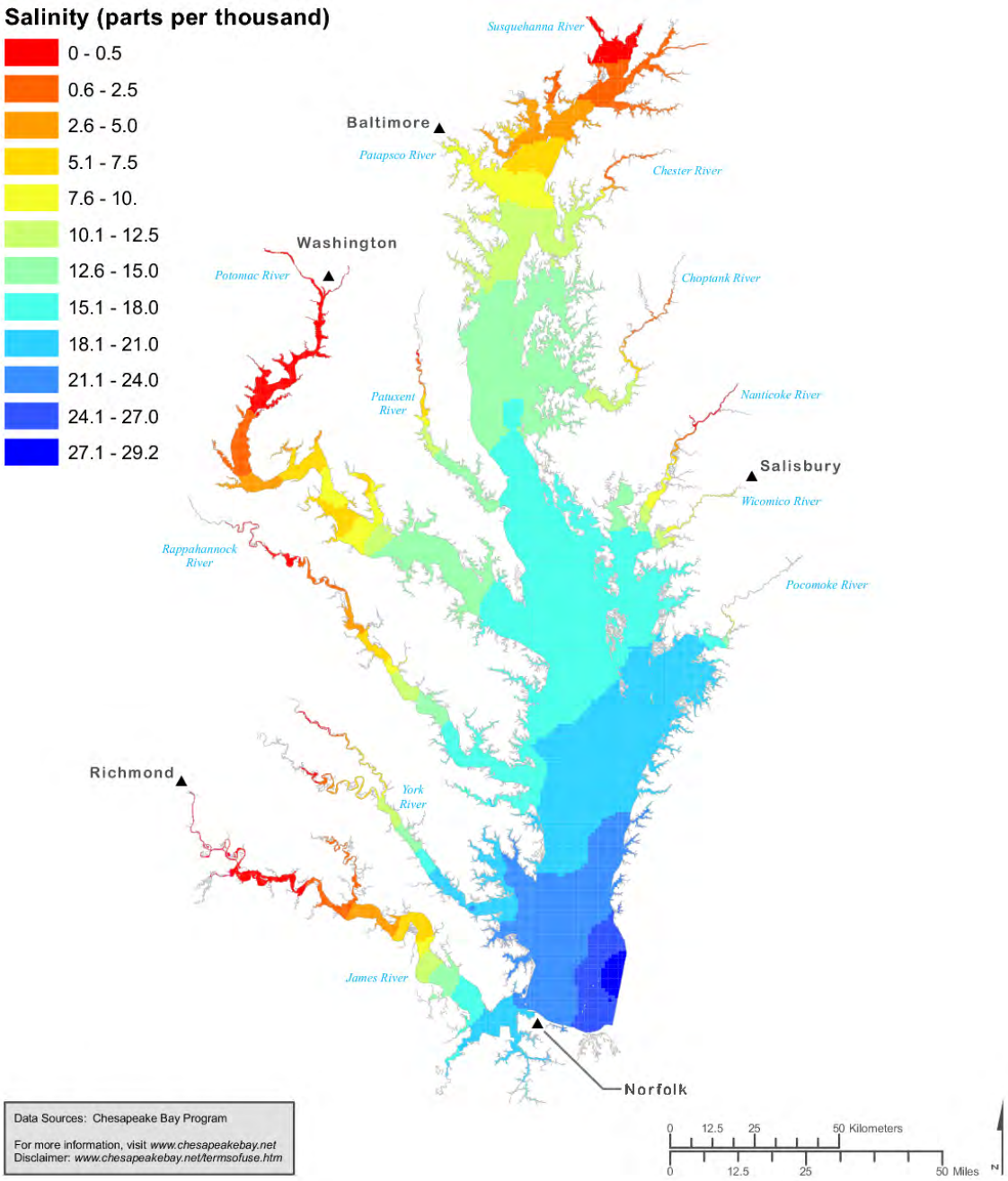
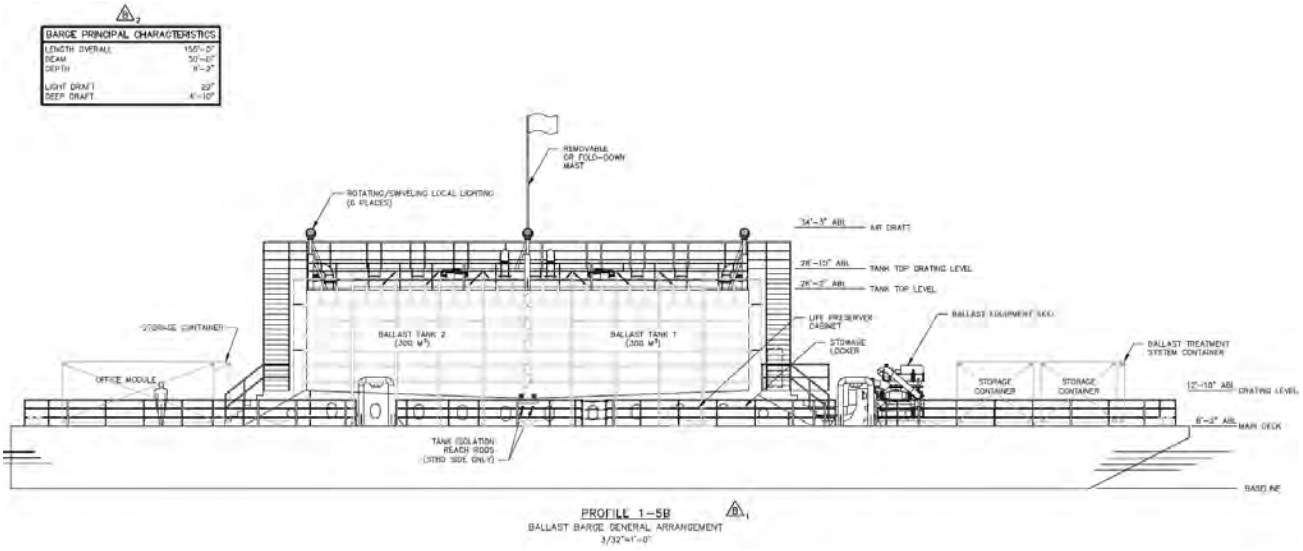


Figure 3. MERC Mobile Test Platform



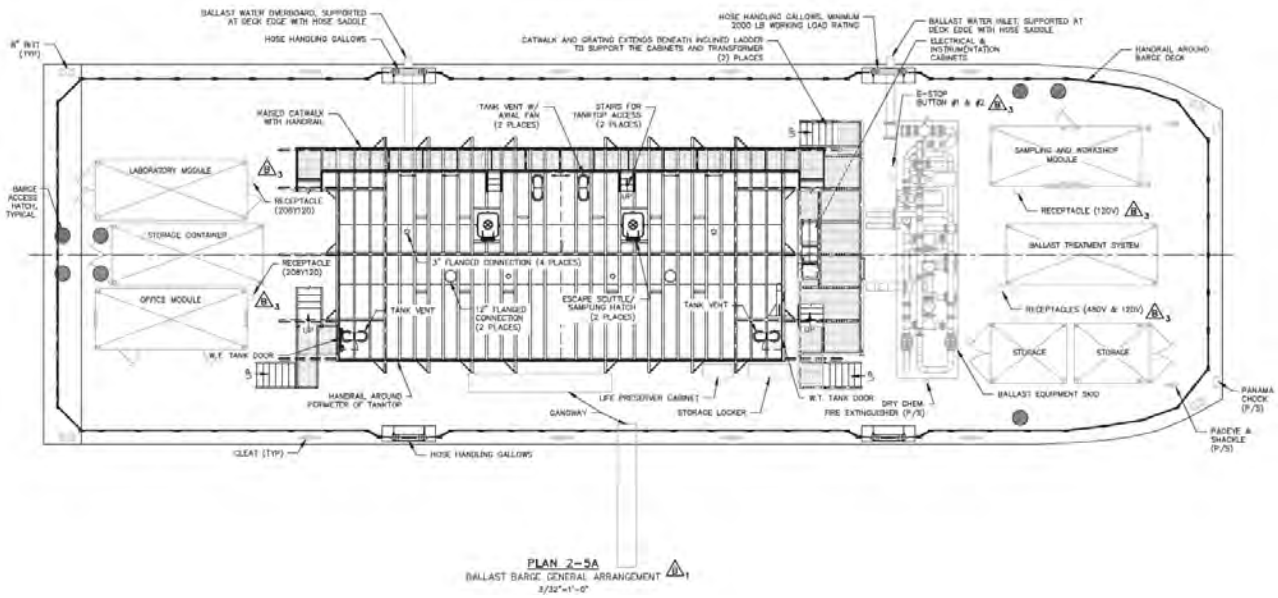


Table 3. Ranges of various physical/chemical and biological parameters in water from Baltimore Harbor, in comparison to ETV and G8 listed challenge conditions

Parameter	ETV	G8	Historic Ranges Port of Baltimore
Temperature (^{o}C)	10 - 35	–	4 - 28
Salinity (psu)	0 - 31	Two salinities, >10	5 - 15
Total Suspended Solids (mg/l)	> 15	> 50	1 - 60
Particulate Organic Carbon	> 1	> 5	0.5 - 6.0
Dissolved Organic Carbon	> 3	> 5	2 - 10
Zooplankton (> 50 μm) / m^3	> 10,000	> 100,000	10,000 - 300,000
Phytoplankton (10 - 50 μm) / ml	> 100	> 1,000	500 - 15,000
Heterotrophic Bacteria cfu / ml	> 1,000	> 10,000	10,000 - 10,000,000

A.7. Quality Objectives and Criteria for Measurement Data

In performing BWTS tests, MERC and all participating laboratory staff will follow the technical and QA procedures specified in this QAPP and will comply with the data quality requirements in the MERC QMP (Section 8.2.3). Data quality objectives (DQOs) have been established as test conditions to ensure that MERC tests provide suitable data for robust evaluations of performance.

A.7.1. Data Quality Objectives

The development of the DQOs follows U.S. EPA's Guidance for the Data Quality Objectives Process (EPA QA/G-4, 2006). DQOs are qualitative and quantitative statements that clarify study objectives, define the appropriate types of data, and specify the tolerable levels of potential decision errors that will be used as the basis for establishing the quality and quantity of data needed to support decisions. DQOs therefore provide the criteria to design a sampling program within cost and resource constraints or technology limitations. DQOs are typically expressed in terms of acceptable uncertainty associated with a point estimate at a desired level of statistical confidence. Acceptance criteria are specifications intended to evaluate the adequacy of one or more existing sources of data as being acceptable to support the project's intended use. Data quality objectives and acceptance criteria vary by analysis type and will be specified in specific test plans. In general, only data that meet or exceed these criteria are deemed valid, thereby ensuring that all data generated is of the highest quality.

A.7.2. Measurement Quality Objectives

Measurement Quality Objectives (MQOs) are a subset of DQOs. MQOs are designed to evaluate and control various phases (sampling, preparation, and analysis) of the measurement process to ensure that total measurement uncertainty is within the range prescribed by the project's DQOs. MQOs define the acceptable quality (data validity) of field and laboratory data for the project. MQOs are defined in terms of the following data quality indicators:

- Accuracy;
- Precision;
- Bias;

- Representativeness;
- Completeness;
- Comparability; and
- Sensitivity

Accuracy and precision are monitored through the analysis of QC samples. Completeness is a calculated value. Sensitivity is monitored through instrument calibration and the determination of method detection limits (MDLs) and reporting limits. Qualitative quality objectives, expressed in terms of comparability and representativeness, are addressed as part of the sampling design.

Accuracy

Accuracy is a measure of the overall agreement between an observed value and an accepted reference value. Accuracy includes a combination of random error (precision) and systematic error (bias) components that are due to sampling and analytical operations. Accuracy is measured through the analysis of matrix spikes and/or laboratory control samples, as and if required by the analytical methods, to determine percent recoveries (%R). A certified reference standard is used if available.

The %R utilizing matrix spikes is calculated as follows:

$$\%R = \frac{(C_s - C_u)}{C_a} \times 100$$

where C_s = measured concentration of spiked sample
C_u = measured concentration of unspiked sample
C_a = actual concentration of spike added

The %R utilizing laboratory control samples is calculated as follows:

$$\%R = \frac{(C_m)}{(C_a)} \times 100$$

where C_m = measured concentration of control sample
C_a = actual concentration of control sample

Accuracy should be assessed using a minimum of 6 determinations over a minimum of 3 concentration levels (e.g. 3 concentrations/ 2 replicates) in a representative pool of sample matrix (preferably the same pool of matrix used to prepare matrix controls). The analyte concentrations tested should be targeted to the same region of the standard curve as the matrix controls. In general, accuracy should be within the range of 70 - 130 percent recovery of exogenous analyte.

Precision

Precision is the degree to which a set of observations or measurements of the same property, obtained under similar conditions, conform to themselves.

The precision of field samples is assessed by the comparison of field duplicates. The relative percent difference (RPD) between the analyte levels measured in the field duplicates is calculated as follows:

$$RPD = \frac{|C_a - C_b|}{0.5(C_a - C_b)} \times 100$$

where C_a = measured concentration of sample
C_b = measured concentration of duplicate sample

The precision of physical parameter readings may be assessed by the comparison of each instrument's calibration readings versus the post check readings. The RPD between the readings is calculated as follows:

$$RPD = \frac{|R_x - R_y|}{0.5(R_x - R_y)} \times 100$$

where R_x = calibration reading
R_y = post check reading

The precision of the laboratory analysis is assessed by the comparison of matrix spikes (MS) and matrix spike duplicates (MSD), if required by the analytical method. The RPD between the analyte levels measured in the MS sample and the MSD sample is calculated as follows:

$$RPD = \frac{|C_{MS} - C_{MSD}|}{0.5(C_{MS} - C_{MSD})} \times 100$$

where C_{MS} = measured concentration of the matrix spike
C_{MSD} = measured concentration of the matrix spike duplicate

For parameters where spiked samples are not practicable to assess laboratory precision, such as live zooplankton, phytoplankton, and bacteria, a comparison of laboratory replicate analyses may be performed in order to calculate the RPD.

For zooplankton samples collected at the MERC Mobile Test Platform, precision is measured by analyzing at least two counting chambers from every sample collected. Precision is quantified by calculating a coefficient of variation (CV) for each sample as follows:

$$CV = \frac{(S)}{(\bar{x})} \times 100$$

For phytoplankton samples collected, at least two out of five treatment discharge samples and at least one out of five control intake or discharge samples (from each set of five test trials) is selected for evaluation of within-sample precision. Precision is measured by the analysis of at least two subsamples by the same phytoplankton taxonomist. In the event that there are fewer than ten total control samples collected during a treatment technology performance evaluation, a minimum of one discharge control sample is chosen for evaluation of within-sample precision.

Bias

Bias is systematic or persistent distortion of a measurement process that causes errors in one direction. Bias may originate from sources such as calibration errors, response factor shifts, unaccounted-for interferences, or chronic sample contamination. The sample itself may generate real or apparent bias caused by a matrix effect or variation in physical properties such as particle size. A bias DQI is a quantitative indicator of the magnitude of systematic error resulting from these effects. Bias can be in the positive (high) or negative (low) direction from the true value and is usually unknown in magnitude.

MERC estimates bias by testing the measurement system result against a specimen with known properties. The most common DQIs for bias are derived from the results of QC samples such as spiked samples, standard reference materials, and various kinds of blanks in the sample stream. Table 4 lists some common MERC QC samples and the components of bias they are intended to measure.

Table 4. QC samples for deriving bias indicators.

Sample Type	Indicator For
Blank spike	Instrument contamination or malfunction, calibration shift
Matrix spike	*plus* effectiveness of sample extraction/digestion procedures
Reference material	Same as matrix spike, but more representative of overall performance when material is similar to matrix examined in the study
Calibration blank	Instrument contamination, calibration shift
Preparation blank	*plus* laboratory contamination
Field blank (equipment/trip)	*plus* field, transportation, and storage contamination

The difference between the measured and expected result is a DQI for bias. For spikes and reference materials, MERC expresses bias as a fractional or percent comparison of the measured result to the expected result.

The percent spike recovery is calculated as follows:

$$\text{percent spike recovery} = \frac{(x_s - x_u)}{x_a} \times 100\%$$

where x_s = measured value of spiked sample
x_u = measured value of unspiked sample
x_a = known amount of spike in sample

A completely unbiased result thus has recovery of 1 (or 100%) and recovery may be greater or less than 1 (100%) depending on whether the result is higher or lower than the known quantity. For blank samples, the actual magnitude of the result is the DQI because the "known" quantity should be zero for a blank.

MERC evaluates taxonomic bias (zooplankton, phytoplankton, bacteria) by comparing whole-sample identifications completed by independent taxonomists. To calculate taxonomic bias, MERC generally randomly selects 10 percent of the samples for recounts and re-identification by a second qualified taxonomist.

Final counts for samples are dependent on the taxonomist. Comparison of counts is quantified by calculation of Relative Percent Difference in Enumeration (RPDE), using the formula:

$$RPDE = \frac{(|x_1 - x_2|)}{\frac{x_1 + x_2}{2}} \times 100\%$$

where x_1 = number of organisms in a sample counted by the first taxonomist
x_2 = the recount by the second taxonomist

Individual samples exceeding 5% are re-examined.

The measure for taxonomic bias is Percent Taxonomic Disagreement (PTD), which is calculated as:

$$PTD = 1 - \frac{(comp_{pos})}{N} \times 100$$

where $comp_{pos}$ = the number of agreements and N = the total number of individuals in the larger of the two counts.

The lower the PTD, the more similar are taxonomic results. Individual samples exceeding 15% are examined for taxonomic areas of substantial disagreement, and the reasons for disagreement investigated. Where re-identification by an independent, outside taxonomist is not practical, percent similarity is calculated. Percent similarity is a measure of similarity between two samples. Values range from 0% for samples with no species in common, to 100% for samples which are identical. It is calculated as follows:

$$PSC = 1 - (0.5 \sum_{i=1}^{K} (|a_i - b_i|) \times 100\%$$

where: *a* and *b* are, for a given species, the relative proportions of the total samples A and B, respectively, which that species represents. The MQO for percent similarity of taxonomic identification is ≥85%. If the MQO is not met, the reasons for the discrepancies between analysts are determined, and the batch of samples with discrepancies may be recounted.

Representativeness

Representativeness, as defined by the American Society for Quality and published in the American National Standards Institute (ANSI) document, ANSI/ASQC E4-1994, *Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs* (ANSI/ASQC, 1994) is: "The measure of the degree to which data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, a process condition, or an environmental condition."

Developing a clear understanding of the "population" that is the subject of the test is the key to assessing representativeness. The characteristics of the population include the subject's identity or class (e.g., the particular property that needs to be measured), the spatial distribution of the property, and in some cases, the temporal characteristics of the property. This definition of representativeness encompasses issues at

both the micro- and macro-scale by addressing both how well measurements taken within a sampling unit reflect that unit and the degree to which measurements from a set of sampling units represent the population of interest.

Representativeness is usually considered a qualitative term. The basic questions to be answered are whether the individual measurements of the characteristics of interest accurately reflect the conditions in the sampling unit, and whether an adequate number of units were measured to reflect the population of interest. It is addressed primarily in the sample design, through the selection of sampling sites and procedures that reflect the test goals and the environment being sampled, i.e., the Mobile Test Platform. A review of the results of quality assessment samples such as field duplicates (collocated samples), splits, or other replicates also is performed. It is ensured in the laboratory through (1) the proper handling, homogenizing, compositing, and storage of samples and (2) analysis within the specified holding times so that the material analyzed reflects the material collected as accurately as possible.

Completeness

Completeness is a measure of the amount of valid data obtained from a measurement system as compared to the amount needed to ensure that the uncertainty or error is within acceptable limits. It is a measure of how well a sampling and analysis design was implemented. It is expressed as follows:

$$\%C = \frac{(M_v)}{(M_p)} \times 100$$

where M_v = number of valid measurements
M_p = number of planned measurements

The goal for data completeness is 100%. Events that may contribute to reduction in measurement completeness include sample container breakage and laboratory equipment failures. Samples are considered invalid if they are contaminated, fail to meet the data quality objectives or other QA protocols, are lost through sample destruction, are incorrectly collected or analyzed, and/or if there is insufficient amount of sample for analysis.

The field and laboratory completeness objectives for each BWTS test are determined during test development and specified in each specific Test Plan. The general completeness criterion for all field measurements and sample collection is 90 percent, but will be influenced by factors mentioned above. If the completeness objectives are not achieved for any particular category of data, the MERC PC will provide documentation why the objective was not met and how the lower percentage impacted the overall study objectives. If the objectives of the study are compromised, re-sampling or re-measurement may be necessary. The respective Senior Researcher/Laboratory Director assures the validity of the analytical measurements reported, and the PC validates the numbers of valid measurements. The completeness criterion for all laboratory measurements is 95 percent, unless specified differently in a Test Plan.

Comparability

Comparability is a qualitative measure of the confidence with which one data set can be compared to another. The key to comparability is consistency of approach, which applies to both the field portion of the sampling and the laboratory analysis of the samples. In the field, it is addressed primarily in sampling design through use of comparable sampling. In the laboratory, comparability is ensured through the use of comparable analytical procedures and ensuring that project staff are trained in the proper application of the procedures. Within-study comparability is assessed through analytical performance (QC samples). The assessment of this DQI determines if analytical results being reported are equivalent to data obtained

from similar analyses. Only comparable data sets can be readily combined. Table 5 presents nine indicators of comparability, and questions are considered related to each.

Table 5. Indicators of comparability.

Indication of Comparability	Related Questions
Samples within data sets should be selected in a similar manner	Sample design: Were the samples selected in a similar manner? Are they equally representative of the population of interest? If samples in one data set were selected using a judgmental sample design, and another data set is based on a statistical design, then combining these data may not be appropriate for some uses.
Data should be temporally and spatially consistent	Sample collection dates: Were samples collected in the same sampling event? Are there temporal factors such as seasonality or holding times that could directly affect interpretation of the data?
	Sample location: Were the samples taken from the same area? Are they representative of the same population spatially? If they are from different areas, how are they expected to be similar? How are they expected to differ?
	Matrix: Were the samples from the same matrix? This relates to how the samples were collected, location of the samples, and when the samples were collected. If matrices are different, are they expected to be related in some way?
Data sets should contain the same set of variables of interest	Variables of Interest: Which variables are of interest and are necessary for grouping or analyzing the data? Were these variables reported for all data sets?
Units in which these variables were measured should be convertible to a common metric	Units: Units should be reported for all data sets. Are the units all convertible to a common metric? For example, some results may be reported in wet weight and some in dry weight, which are not directly comparable without additional information
Field collection methods should be similar	Field methods: What instrument was used and which procedure was followed? Were single or composited samples collected?
	Sample handling: Some samples require special handling such as preservatives or special containers. Differences in sample handling may cause variations in the results, which may affect comparability. Were the samples filtered or unfiltered? Are there chain-of-custody forms available for all samples?
Similar sample Preparation methods should be used	Laboratory: Was the same laboratory used for all analyses? The use of routine methods and procedures simplify the issues of comparability because the same standards should be met. In addition, this will increase confidence in the comparability of methods used.
	Sample preparation: Was the same sample preparation used for all samples? If not, are the sample preparation methods comparable?
Similar procedures and quality assurance should be used to collect and analyze samples for all data sets	Analytical method: Was the same analytical method used for all samples? If not, are any of the analytical methods comparable? The use of routine methods simplifies the determination of comparability because all laboratories used the same standardized procedures and reporting parameters. However, when reviewing the analytical methods, consideration must also be given to options that may be available within the method. Although the analytical method may be the same, options such as matrix or concentration level will affect results reported.
	Analytical method options: If the analytical methods are comparable, were the same options within each method chosen? The options available within each method must also be checked because the same analytical method using different options may produce very different results.
Measuring devices used for both data sets should have approximately similar detection levels	Detection or quantitation level: Are non-detects generally reported at the same level? Are the detection or quantitation levels acceptable for use in decision making? Combining data sets having different detection or quantitation levels leads to difficulties in analytical interpretations.
	Quality control of data entry, storage, transfer, and retrieval: Were results reported into the database in a consistent manner? Have all data sets been checked for completeness?

Indication of Comparability	Related Questions
Rules for excluding certain types of observations should be similar for all data sets	Qualification and/or validation of data: What criteria were used to qualify or validate the data? If criteria were not consistent across data sets, the same qualifications may have different meanings. What QA and QC information is available from the laboratories?

Sensitivity

Sensitivity is the capability of a test method or instrument to discriminate between measurement responses representing different levels (e.g., concentrations) of a variable of interest. Both the precision of the instrument and the slope of the calibration curve limit sensitivity.. If two methods have equal precision, the one having a steeper calibration curve will be the more sensitive. Sensitivity can also be evaluated from the standard deviation of replicate analyses at any concentration level. Sensitivity is addressed primarily through the selection of appropriate analytical methods, equipment, and instrumentation.

The sensitivity indicators of primary interest to MERC are indicators that relate to limits of detection. The detection limit (DL) is a concept concerning the capability of an analytical method to distinguish samples that do not contain a specific analyte or biological variable from samples that contain low concentrations. DLs vary by variable and by matrix.

The sensitivity of microbial analyses utilizing an IDEXX® analytical method is reported in the product literature for each type of analysis (i.e., *E. coli*, *Enterococcus spp.*, total heterotrophic bacteria).

MERC uses two sensitivity indicators to define detectability for chemical analyses: method detection limit (MDL) and practical quantification limit (PQL) or reporting limit (RL).

The MDL is the minimum value which the instrument can discern above background but no certainty to the accuracy of the measured value. MDLs facilitate the determination of whether a single observation represents a true signal as opposed to noise. Two approaches have generally been recommended to determine MDLs, “Single Concentration Designs" and "Calibration Designs". The single concentration design assumes variability at a certain concentration to be equal to the variability at the true MDL. The calibration method utilizes prediction intervals to model MDL variance with concentration (Hubaux and Vos, 1970).

MERC utilizes the single concentration design estimator to determine MDL, which is recommended by the US EPA:

1) Measurements are taken on at least seven samples of the prepared solution. Results are tabulated and the standard deviation of the data set is taken:

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\right]$$

2) Using the degrees of freedom from the data set and the appropriate confidence level (usually 1%), the critical t-value is looked up using reference tables:

$$t_{\nu,\alpha} = (\text{look - up value})$$

3) The MDL is computed as the product of the standard deviation and the critical t-value:

$$MDL = t_{\nu,\alpha} \times s$$

4) To test for uniform variance, another solution is spiked with a slightly different concentration of the analyte. Measurements are taken on at least 7 samples of this new solution and results are tabulated as with the first solution.
5) An F-test for two sample variance is performed on the two data sets to ensure that the difference between the variances are "statistically insignificant" ($F_{stat} < F_{crit}$).
6) If determined insignificant, the procedure continues by pooling the two sample variances as follows:

$$s^2_{pooled} = \frac{\nu_1 s_1^2 + \nu_2 s_2^2}{\nu_1 + \nu_2}$$

7) A new critical t-value is looked up using the new amount of degrees of freedom:

$$t_{\nu 1+\nu 2,\alpha} = (\text{look - up value})$$

8) The MDL is computed as the product of the new critical t-value and the pooled standard deviation:

$$MDL = t_{\nu 1+\nu 2,\alpha} \times s_{pooled}$$

The PQL or RL is the minimum value that can be reported with confidence, i.e., some multiple of the MDL. The requirements for quantification are more stringent than for detection. A limit of quantification (LOQ) is generally defined at 5 to 10 times the standard deviation of the noise (or blank) signal. In practice, PQLs are defined at the lowest concentration in the calibration curve. In this method, the calibration statistics (e.g., correlation coefficient, confidence intervals, RSD of response factor) ensure that the PQL represents the same precision and accuracy as other data reported for the analyte.

MERC reports sample data measured below the MDL as ND or non-detect. Sample data measured ≥ MDL but ≤ PQL or RL is reported as estimated data. Sample data measured above the PQL or RL is reported as reliable data unless otherwise qualified per the specific sample analysis.

A.8. Special Training/Certification

The MERC QMP (Section 4.0) requires that staff members have the knowledge, skill, and any professional certifications needed to perform their MERC assignments and that quality management responsibilities and requirements are understood at every stage of project implementation throughout MERC (Table 6). The PC is responsible for identifying worker certification needs and ensuring that all team members are adequately trained.

A.8.1. General Training Requirements

The general approach to training is to utilize a combination of clearly documented guidance material and experience from existing or previous staff and supplement these with formal trainings as needed. Individuals implementing MERC tests must receive, at a minimum, orientation to the project's purpose, scope, and methods of implementation. This orientation is the responsibility of the PC.

The MERC PC is responsible for determining specific training and certification needs, and for ensuring that any required training is documented. The MERC PC and Senior Scientists are responsible to ensure that all field and laboratory personnel receive orientation to applicable policies, procedures, requirements, and their scope of application.

All field and laboratory staffs are to be trained, at a minimum, in the methods described in all applicable SOPs. Initial training of field personnel in activities such as instrument calibration, safety, required documentation, sampling methods, sample handling, and safety is generally performed by the MERC PC. The PC also may conduct a field orientation or pre-deployment practice sessions. MERC Senior Scientists are responsible for ensuring that technicians under their supervision possess and maintain adequate proficiency, expertise, and knowledge in their respective work disciplines. Each laboratory technician and analyst must complete an initial demonstration of capability before processing or analyzing samples for MERC tests. Information on laboratory staff competence should be provided in each lab's Quality Management and/or Quality Assurance Plan. Current copies of the laboratory's QA Plan and attendant method specific SOPs should be on file with the MERC PC during the duration of laboratory use.

Field and laboratory staffs are assessed on an ongoing basis by their direct supervisor and the MERC PC to ensure all technical staff are performing activities in accordance with SOPs, the QAPP, and the Test Plan. Experienced field and laboratory staff will continue to work with all new staff during sampling and analytical activities until the new staff member exhibits proficiency in the field and/or laboratory, as determined by the trainer's observations.

At least annually, field and laboratory technicians and analysts must demonstrate continued proficiency for the tasks that they are performing. The procedures used to ensure that staff training is current and documented are defined in the QMP and SOPs.

Documentation of training related to technology testing, data analysis and reporting is maintained for all MERC and MERC Partner technical staff in training files at their respective locations. The MERC QA Manager may verify the presence of appropriate training records prior to the start of testing.

A.8.2. Test-Specific Training

Manufacturers of BWT technologies will be required to train the MERC technical staff prior to the start of testing. MERC will document this training with a consent form, signed by the vendor, which states which MERC technical staff have been trained to use their technologies and can train other staff. In the event that other staff members are required to use the technologies, they will be trained by the operators that were trained by the vendors.

Table 6: Training

Specialized Training/Certification	Field Staff	Program Coordinator	Laboratory Staff	Senior Researchers	QA Manager
Safety training	X	X	X	X	X
Sampling techniques	X	X			X
Instrument calibration and QC activities for field measurements	X	X			X
Instrument calibration and QC activities for laboratory measurements			X	X	X
QA principles		X		X	X
Chain of Custody procedures for samples and data	X	X	X	X	X
Field Measurement Method Training	X	X			X
Lab Analytical Method Training			X	X	X
Specific BWTS Operation Training	X	X			X

A.9. Documentation and Records

The documents for each MERC BWTS test will include the QAPP, the Test Plan, vendor instructions, reference methods, the verification report, verification statements, and audit reports. The project records will include certificates of analysis (COA), chain-of-custody forms, laboratory record books (LRB), data collection forms, electronic files (both raw data and spreadsheets), and QA audit files. All of these documents and records will be maintained by the MERC PC during the tests and will be transferred to permanent storage at the conclusion of the verification tests. Section 6 of the MERC QMP further details the data recording practices and responsibilities.

All data generated during the conduct of this project will be recorded directly, promptly, and legibly in ink. All data entries will be dated on the date of entry and signed or initialed by the person entering the data. Any changes in entries will be made so as not to obscure the original entry, will be dated and signed or initialed at the time of the change and shall indicate the reason for the change. Project-specific data forms will be developed prior to testing to ensure that all critical information is documented in real time. The draft forms will be provided to the MERC QA Manager for review prior to use so that appropriate changes, if any, can be made.

B. MEASUREMENT AND DATA ACQUISITION

B.1. Experimental Design

The goal of MERC evaluations of BWTS is to verify the biological treatment performance according to established U.S. and international protocols and specified challenge conditions identified in an approved Test Plan. MERC BWTS tests are designed to:

- Provide a comprehensive set of water quality and biological challenge conditions against which treatment effectiveness can be quantitatively evaluated.
- Develop adequate data to document system performance against the verification factors.

As noted previously, MERC protocols, challenge conditions and testing infrastructure are based on the essential features of the G8 guidelines and ETV protocols for the verification of BWT technologies. To that end, all BWTS will be evaluated on the following:

- Biological treatment efficacy
- Operation and maintenance
- Reliability
- Cost factors
- Environmental acceptability
- Safety

Biological treatment efficacy is measured in terms of the concentration of selected organism size classes in the challenge water and the treated discharge. Operation and maintenance includes the labor, expertise, equipment, and consumables required to operate the system to achieve the stated performance goals and objectives. Reliability is a statistical measure of the number of failures (either qualitative or quantitative) per known quantity of test cycles. Cost factors include only those factors that can be verified, such as labor hours to operate and maintain the system, expendable material, such as filter cartridges, and pounds or gallons of chemicals consumed by the treatment system. Environmental acceptability assesses ballast water quality following treatment for factors such as whether the treated water meets acceptable water quality characteristics. Safety factors include any treatment-specific considerations that may pose a threat to the safety of the operator or shipboard operations.

Another basic premise in the design for the test design is that BWTS are designed to function effectively in the full range of water quality characteristics that will be encountered under shipboard operational conditions. By challenging the treatment systems with these conditions, it is assumed treatment will be effective under less demanding conditions. The measurement methods for evaluating the status of the challenge water quality conditions are described in Section B.4.

Biological efficacy will be evaluated as function of a system's ability to kill or remove organisms that are naturally occurring and represent the more robust ambient populations at the test site. A minimum total input concentration of living organisms, by size class, as describe in A.6.1. Upon intake, challenge water conditions will be determined for each test cycle prior to water entry into control tanks, Samples of treated water are collected immediately downstream of the treatment system. At the end of the hold time specified in the test plan, samples are collected upon discharge from both the treated and control holding tanks. Treatment tests will evaluate equipment at operational flow rates defined by the vendor's O&M manual. A minimum of 400 m^3 per tank, divided equally into control and treated tanks shall be processed in each test cycle.

B.2. Sampling Method Requirements

As described in Section A.6, MERC BWTS testing takes place on MERC's Mobile Test Platform, which may be located at three sites, Baltimore Harbor, Baltimore, MD; Norfolk, VA; or the Anacostia River in Washington, DC. Ambient conditions at each site are employed as the physical/chemical challenge conditions, although certain parameters may be augmented to meet IMO or ETV requirements. These will be detailed in the Test Plan. Biological challenge conditions are also ambient but can be enhanced, if required.

Flow control valves and system logic assure that sample flow rates are equivalent and proportional to intake and discharge flow rates throughout each operation. Flow rates are recorded every 15 seconds

during the test trials by automated meters located on the control track, treatment track, and on the discharge line. Pressure readings are also recorded every 15 seconds throughout the facility.
Samples for water chemistry and water quality analysis can be collected during intake, tank retention and discharge and will be specified in the test plan. The water chemistry and water quality parameters that are measured are also specific to the treatment technology being evaluated, and will be detailed in each test plan.

In addition, temperature, dissolved oxygen, turbidity, chlorophyll, conductivity, and salinity are measured regularly throughout the retention period by two identical multi-parameter sondes (calibrated according to manufactures specifications) placed, one each, into the central mid-water of the control and test tanks. A calibrated, hand-held sonde may also be used to measure temperature, salinity, pH, turbidity, conductivity, chlorophyll and dissolved oxygen from the control sample collection tubs, the pre-treatment sample collection tubs, and the post-treatment sample collection tubs during intake. These parameters may also be measured during discharge from the control and treatment sample collection tubs. The specifics of these measurements will be detailed in the test plan.

Continuous (time-integrated) sampling occurs at a fixed rate during uptake and discharge to assure the highest statistical confidence in results (see Miller et al., 2011). Two sampling assemblies reside in a modified, bow-mounted 20-foot shipping container with garage-doors, lights, outlets, and other alterations. One sampling assembly is dedicated to each tank and pump/pipe array. Sample flow is available in every ballast-operation mode. Pressure at sample-points is similar to pump-discharge pressure except post-vendor which may be 6 or 7 psi lower. Ballast system operator will maintain sample-point pressure sufficient for sampling rates of between 3 m^3/hour and 10 m^3/hour without need for additional sampling pumps. Sampling rates are monitored via precise magnetic flow-meters logging their data to a process-control system. Rate adjustments occur via diaphragm valves allowing fine-grain adjustments. The >50 µm biota fraction are removed from sampling flow via 35 µm square-pore water-suspended net-collectors. Continuous (time-integrated) sampling occurs during uptake or discharge events, as required by ETV and IMO protocol. Wastewater from this process is directed to overboard drains. Whole-water samples (not sieved) are taken for the 10 µm to 50 µm biota and water quality assays. They also are time-integrated.

All samples collected to quantify live organisms or water quality will be taken by inline sampling of water during the entire filling or discharge of water from the tanks through sample ports located on appropriate filling or discharge pipes. All sample ports include a valve and sample tube with a 90^o bend towards the direction of flow, placed in the center of the piping system (based on the design developed and validated by the US Naval Research Laboratory, Key West Florida, see ETV protocols).

B.3. Sample Handling and Custody Requirements

Each sample from the Mobile Test Platform will be labeled with a unique sample identifier code to ensure proper identification in the field and/or tracking in the laboratory. These codes are used for the sample containers, field and laboratory data sheets, logbooks, chain of custody forms, and database entries. Sample labels are prepared and placed on sample collection containers prior to sample preparation/collection.

All samples will be handled, prepared, transported and stored in a manner so as to minimize bulk loss, analyte loss, contamination or biological degradation using methods as described in the specific Test Plan and/or according to the procedures presented in specific SOPs for the method in question.

The individual collecting the sample is personally responsible for the handling and custody of the sample until it is transferred to the individual responsible for analyzing the sample. The MERC PC determines whether proper custody procedures are followed during the field work and decides if additional samples are required due to improper sample handling.

Chain-of-custody procedures are strictly followed for all samples that are transported from the Mobile Test Platform to MERC Partner and contracted analytical laboratories so that the possession of a sample from the time of its collection until the time of its analysis is traceable and documentable. These procedures not only guarantee the integrity of a sample (i.e., that it was properly prepared, preserved and/or handled leading up to analysis), but also alleviate the possibility of sample mix-ups and/or extraneous contamination.

A person will have custody of a sample when the samples are:

- in their physical possession;
- in their view after being in their possession;
- in their personal possession and secured to present tampering;
- in a restricted area accessible only to authorized personnel;
- the person is one of the authorized personnel.

Field custody documentation will consist of both field log books and chain of custody forms. Completed chain-of-custody forms will be required for all samples to be analyzed. Chain-of-custody forms will be filled-out by MERC field sampling staff during the sample collection events. The chain-of-custody form will track sample release from the sampling location to the analysis laboratory. The chain-of-custody form will contain each sample's:

- unique identification number;
- sample date and time;
- sample description;
- sample type
- sample preservation (if any);
- analyses required.

The original chain-of-custody form will remain with the samples at all times. Each form will be signed by the person relinquishing samples once that person has verified that the chain-of-custody form is accurate. Copies will be made prior to shipment for separate field documentation and retained by the individual relinquishing the sample.

Upon arrival at the analysis laboratory, chain-of-custody forms will be signed by the person receiving the samples (if different from the sample collector) once that person has verified that all samples identified on the chain-of-custody forms are present.

Laboratory sample custody will be performed in accordance with the laboratory's Quality Assurance Manual and SOPs and will be consistent with the guidelines set forth in this section of the QAPP. The procedures should include but not be limited to documenting the following information:

- presence or absence of chain-of-custody forms,
- presence of absence of bills of lading
- presence or absence of custody seals on shipping and/or sample containers and their conditions,
- presence or absence of sample labels,

- sample label id numbers if not recorded on the chain-of-custody record(s) or packing list(s),
- condition of the shipping container,
- condition of the sample bottles,
- verification of agreement or nonagreement of information on receiving documents,
- resolution of problems or discrepancies

After samples are received, they are placed in secure storage (e.g., locked refrigerators). The laboratory will have written SOPs which specifically include descriptions of all storage areas for samples in the laboratory, and steps taken to prevent sample contamination. Only authorized personnel will have access or keys to secure storage areas. Each laboratory also will have written SOPs for tracking the work performed on any particular sample. Sample receipt, sample storage, sample transfers, sample preparations, sample analyses, instrument calibration and other QA/QC activities will be documented. All samples remaining after successful completion of analyses will be disposed of properly, in accordance with all applicable regulations.

All relevant MERC senior personnel are responsible for ensuring that the chain-of-custody forms are correctly filled out at the time of changes to sample custody, and sample handling and storage. They are also responsible for maintaining the forms with the test records.

B.4. Analytical Method Requirements

The analytical methods used by MERC and any contract laboratories for BWTS tests are in accordance with procedures currently approved or accepted by the USEPA. A summary of the analytical methods and approximate data turnaround times are described in Table 7. A more detailed description of the analytical equipment and instrumentation required for each analysis is included in the individual field or laboratory SOP for the methods.

Table 7. Analytical methods and reference limits for core parameters

Category	Core Parameter	Reporting Units	Analytical Method	Acceptable Range for Initiating Testing
Physical/ Chemical	Salinity	PSU	SOP - Ballast System Instrumentation	0 – 36
	Dissolved Organic Carbon (DOC)	mg/l	SOP – Water Quality Analyses SOP – Challenge Water Modification	Can be artificially augmented, dependent on test plan
	Particulate Organic Carbon (POC)	mg/l	SOP – Water Quality Analyses SOP – Challenge Water Modification	Can be artificially augmented, dependent on test plan
	Total Suspended Solids (TSS)	mg/l	SOP – Water Quality Analyses SOP – Challenge Water Modification	Can be artificially augmented, dependent on test plan
	Dissolved Oxygen	mg/l	SOP - Ballast System Instrumentation	3 - 15
	pH	--	SOP - Ballast System Instrumentation	7 - 9
	Temperature	°C	SOP - Ballast System Instrumentation	4 - 30
	Water Flow Rate	m^3/hr	SOP - Ballast System	100 - 350

Category	Core Parameter	Reporting Units	Analytical Method	Acceptable Range for Initiating Testing
			Instrumentation	
Biological	Zooplankton, live organisms ≥ 50 µm in size	Live Organisms/ m^3	SOP – Live Organisms >50 Microns	Organisms ≥ 50 µm in minimum dimension should be present in a total density of >75,000 to live individuals per m3, and should consist of at least 5 species from at least 3 different phyla/divisions. Culture organisms can be added.
	Protists, live organisms 10 - 50 µm in size	Live Individuals /ml	SOP – Live Organisms 10 - 50 Microns	Organisms 10 - 50 µm in minimum dimension should be present in a total density of not less than 750 cells per mL, and consist of at least 5 species from at least 3 different phyla. Culture organisms can be added.
	Bacteria	Viable bacteria/ml	SOP- Live Bacteria and Indicator Pathogens	Heterotrophic bacteria should be present in a density of at least 1,000 per ml.

When problems occur during the analytical process, a corrective action is implemented. The corrective action should identify the source of the problem and eliminate it. Staff communicates corrective actions to management to determine if additional corrective actions are necessary. The senior researcher of each lab has the primary responsibility for responding to failure of analytical systems. Solutions, which are consistent with the measurement objectives, will be reached in consultation with the MERC QA Manager.

Failures in field and laboratory measurement systems involve, but are not limited to, such things as instrument malfunctions, failures in calibration, sample jar breakage, blank contamination, and quality control samples outside of defined limits In many cases, field staff or lab analysts are able to correct the problem. If the problem is resolvable by MERC field staff or lab analysts, then they document the problem in their field data sheet or laboratory record and complete the analysis. If the problem is not resolvable, then it must be conveyed to the respective senior researcher, who makes the determination if the problem compromised the sample analysis and should therefore results not be reported. The nature and disposition of the unresolved problem needs to be documented in the data report that is sent to the MERC PC and QA Manager.

Unused raw sample volume, sample extract and sample digestates are disposed of properly in accordance with each laboratory's waste management procedures. Disposal of unused raw sample for routine analysis will occur when the analysis is complete and verified to be accurate or when holding times are exceeded, whichever is less.

B.4.1. Viable Organisms >50µm in size

The sampling system consists of two sets of paired tanks, each designed to accommodate a 37µm um (50µm diagonally) mesh plankton net used to collect the >50µm size fraction. One pair handles water from the treated ballast tank and the other pair handles water from the untreated (control) tank. The paired sampling tank/net arrangement allows for the residual from the cod-end of one net from each pair to be processed for examination while filtration continues via the other net, thereby avoiding clogging. In this way unimpaired filtration back and forth between each pair of nets continues until a total of 3 to 7 m^3 has

been processed from each of the treated and untreated ballast tanks respectively. The sampling tanks are designed to allow complete immersion of each net during the filtration process, thereby minimizing trauma to filtered organisms. At the end of each trial (after five days), the control and treated ballast tanks are drained and processed as described above, with the treated sample undergoing a second pass through the UV irradiation unit (but not the filter) before sampling.

The proportion and total concentration of live versus dead organisms will be determined using standard movement and response to stimuli techniques and this live/dead analysis will take place within two hours of collecting the individual samples. Three m^3 is the volume collected for control water upon filling and discharge from test tanks (high numbers of live organisms) and 7 m^3 is collected for treated water on discharge after five days (presumably very few live organisms). Depending on concentrations, quantification of zooplankton in initial samples (upon ballasting) and control samples may require analysis of sub-samples and extrapolation to the entire 3 m^3. Zooplankton samples will then also be fixed with buffered, 10% formalin in 500ml Nalgene bottles and shipped to the Smithsonian Environmental Research Center (SERC) for additional taxonomic evaluation. Total counts and general taxonomic classification will be conducted under a dissecting microscope at 25X, except for some taxa, which will be removed and identified using a compound microscope. Larval forms of invertebrates will be identified to higher taxonomic levels such as order (e.g., Decapoda) suborder (e.g., Balanomorpha) or class (e.g., Bivalvia). Adults will be identified to species in most cases.

B.4.2. Viable Organism 10 - 50 µm in size

A 75 L integrated sample will be collected as an unfiltered split sample in parallel with the >50µm fraction. This sample will be the source water for all other analyses including the 10-50µm fraction. 2 l from this integrated sample will be subject to detailed analysis and counting. Determination of concentrations of viable organisms in this size class will be made using four distinct methods (described below).

Analytical methods are described below and in SOPs. All live unfiltered samples will begin to be examined within three hours of collection on the MERC Mobile Test Platform or nearby partner laboratories, which combine state of the art facilities and depth of experience. Any preserved samples are also transported to MERC partner laboratories for further analyses and taxonomic identification.
One 250 ml sub-sample will be stained using a combination of CMFDA (5-chloromethylfluorescein diacetate) and FDA (fluorescein diacetate) as a selective live/viable indicator. Samples stained with CMFDA+FDA, are incubated and observed on a Sedgewick Rafter slide using a Leitz Laborlux S modified for epifluorescence. Cells are scored as live when showing strong fluorescence signature under excitation (some cells also showed motility). This approach has been validated for use in the Chesapeake Bay (Steinberg et al., 2011) and provides the data for comparison to the IMO D2 and USCG Phase 1 discharge standards.

As supporting information, three other sub-samples are analyzed. A second 250 ml is collected from the initial 2 l and fixed with standard Lugol's solution in a amber Nalgene bottles to estimate total cell abundances (both live and dead) and for species identification under an inverted compound microscope using grid settlement columns and phase contrast lighting. A third sub-sample is filtered (Whatman GF/F

0.7 µm pore, 2.5 cm diameter membrane) and frozen (-80^{o}C) until analysis of total active chlorophyll-a by the CBL/UMCES Nutrient Analytical Services Laboratory using US EPA Methods 445.0 for extractive/fluorometric techniques (see Appendix D). Finally, a fourth sub-sample is used to provide a qualitative measure of phytoplankton growth potential by determining chlorophyll levels after samples are allowed to regrow under favorable conditions. Algae specific vitamins, minerals, and nutrients (Guillard 1975, F/2 formulation) are added to three 250 ml sub-samples from both control and treated water collected upon discharge after the five-day hold time and are placed in a standard algal culture light-dark regimen for five to six days, prior to extractive chlorophyll-a analysis (described above).

B.4.3. Viable Bacteria and Indicator Pathogens

An unfiltered 1 l sample of water sub-sampled from an integrated 75 l sample will be analyzed to determine concentrations of total heterotrophic bacteria and three specific indicator pathogens, *E. coli*, intestinal *Enterococci*, and toxigenic *Vibrio cholerae*.

Total heterotrophic bacteria will be enumerated by spread plate method using marine agar (MA) or R2A agar for freshwater according to *Standards Methods for the Examination of Water and Wastewater* (21st edition, 2005). The presence and abundance of *E. coli* and intestinal *Enterococci* is determined using a commercially available chromogenic substrate method (IDEXX Laboratories, Inc.; Noble et al. 2003) and 10 ml and 100 ml water sample aliquots. Additionally, concentrations of culturable *E. coli* and intestinal *Enterococci* are determined using a standard USEPA method, namely, membrane filtration on mTEC agar (*E. coli*) (1 ml, 10 ml and 100 ml) and mEA agar (*Enterococcus*) (10 ml and 100 ml). Finally, the abundance of total and toxigenic *V. cholerae* will be determined by filtration and selection on TCBS agar and enumerated using species-specific RNA colony blot (500 µl to 1 ml) and *ctxA* DNA colony blot (1-10 ml). Viable toxigenic cells of *V. cholerae* are assayed with a commercial DFA kit specific for serogroup O1 (New Horizons Diagnostics) using monoclonal antibodies tagged with fluorescein isothiocyanate (FITC) (Hasan et al. 1994).

B.4.4. Quantifying Physical Conditions

Temperature, salinity, dissolved oxygen, chlorophyll fluorescence, turbidity and pH will be measured every 15 minutes during the test trials by two identical multi-parameter probes (calibrated before each trial according to manufactures specification) placed, one each, into the control and test tanks. A third hand-held instrument will be used to measure temperature, salinity, and dissolved oxygen of water in each uptake or discharge as they occur.

Initial inline samples (three replicates, 500 ml - 2 l each) of ballast water during the filling of the control and test tanks will also be collected, filtered, and analyzed for the water quality parameters of particulate organic carbon (POC), dissolved organic carbon (DOC), and total suspended solids (TSS). See SOPs for details.

B.4.5. Treatment Toxicity

Even if the treatment system systems does not employ an active substance, MERC will conduct at least one set of toxicity tests as part of one trial at each location/salinity (Baltimore and Norfolk. The testing is designed to meet IMO G9 requirements and uses test methods and species employed by the EPA for Whole Effluent Toxicity (WET) testing of effluents (EPA 2002 and ASTM 2006).

A fish, an invertebrate and a plant (algae) will be used in all ballast discharge tests. Because the Chesapeake Bay is a mesohaline aquatic environment with salinities ranging from 5 to 25 psu, estuarine organisms will be used in these tests. The fish species used in the test will be the sheepshead minnow (*Cyprinodon variegatus*), invertebrate will be a mysid (*Americamysis bahia*; formerly *Mysidopsis bahia*) and the algal species will be *Isochrysis galbana*, all listed as estuarine test species in EPA's Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms (EPA, 2002).

B.5. Quality Control Requirements

There is potential for variability in any sample collection, analysis, or measurement activity. Field variability generally contributes more than laboratory variability. Total study error can result from between sampling unit variability (influenced by design error, inherent spatial variability) and within-sampling-unit variability (due to sampling, analytical and data manipulation). MERC has implemented Quality Management System, documented in the MERC Quality Management Plan, which describes quality assurance and quality control (QA/QC) measures that are applied to all aspects of MERC BWTS testing.

Quality assurance measures undertaken to assure the reliability of the data collected include:

- Duplicate sampling to ensure sample representativeness with respect to sampling and handling procedures. The acceptable range of relative percent difference between a sample and its duplicate is 20% unless specified otherwise in a specific test plan. Data falling outside this range is invalid.
- Replicate analysis to ensure sample representativeness with respect to sample processing and analysis. The acceptable range of relative standard deviation among replicate analyzes is 10%. Data falling outside this range is invalid.
- Calibration and maintenance procedures, schedules, and standards (if applicable) for all equipment used in the test. For example, temperatures of refrigerators and incubators are verified with independent thermometers on a regular basis. Calibration of pH probes is performed using appropriate standard solution ranges, etc.

Quality control measures are actions to assure that defined standards are met in the analysis of data. These measures are analyte or method specific and are defined within the relevant SOP. MERC quality control measures include:

- Method blanks to ensure the workspace, handling procedures, and reagents are free from contamination.
- Positive and negative controls establish that the method is working as designed.
- Matrix spikes, to determine percent recoveries of specific analytical methods. Known quantities of the analyte are added to the sample. After subtracting out the background level naturally present in the unspiked sample, a percent recovery can be calculated by dividing the measured value by the known spiked value.

A summary of QA/QC check samples and information they provide are noted in Table 8. Procedures and formulas for calculating applicable QC statistics are described in Section A.7 of this QAPP.

Table 8. QA/QC check samples

QC Check	Information Provided
BLANKS	
Bottle blank	Cleanliness
Field blank	Transport, storage, and field handling bias
Reagent blank	Contaminated reagent
Rinsate or equipment blank	Contaminated equipment
Method blank	Response of an entire laboratory analytical system
SPIKES	
Matrix Spike	Analytical (preparation + analysis) bias
Matrix spike replicate	Analytical bias and precision
Analysis matrix spike	Instrument bias
Surrogate spike	Analytical bias
CALIBRATION CHECK SAMPLES	
Zero check	Calibration drift and memory effect
Span check	Calibration drift and memory effect
Mid-range check	Calibration drift and memory effect
REPLICATES, SPLITS, ETC.	
Field collected samples	Sampling + measurement precision
Field replicates	Precision of all steps after acquisition
Field splits	Shipping + inter-laboratory precision
Laboratory splits	Inter-laboratory precision
Laboratory replicates	Analytical precision
Analysis replicates	Instrument precision

B.6. Instrument/Equipment Testing, Inspection, and Maintenance

All field and laboratory instruments and equipment used in the BWTS tests will be inspected and maintained in accordance to the manufacturers' recommendations, instruction manuals, or the laboratory SOPs of the analysis laboratory. Field equipment refers to items used for on-site monitoring and testing, whereas laboratory equipment refers to items used in the laboratory in support of data collection (e.g., refrigerators). Laboratory instruments are items used for sample analysis.

The MERC Facility Manager is responsible for ensuring regular cleaning, inspection, and maintenance of field equipment. All equipment should be visually inspected daily for damage or dirt, and repaired or cleaned if needed before use. If meters are stored for long periods (> 1 week) without being used, they will be calibrated and inspected at least weekly to keep them in good working order.

Laboratory instrument/equipment maintenance is the responsibility of the MERC senior researchers in their respective laboratories. Any sub-contracted laboratories will follow the testing, inspection and maintenance procedures as stated in the respective laboratory's QAPP and SOPs.

All routine maintenance and non-routine repairs will be documented in a bound logbook and kept with the instrument/equipment. The information recorded will include analyst initials, date maintenance was performed, a description of the maintenance activity, and (if the maintenance was performed in response to a specific instrument performance problem) the result of retesting to demonstrate that the instrument

performance had been returned to acceptable standards prior to reuse. An inventory of critical spare parts is maintained at the Mobile Test Platform and the laboratories to ensure rapid response to issues.

B.7. Instrument Calibration and Frequency

Calibration procedures for field equipment will follow manufacturer instructions. Laboratory instrument calibration will follow manufacturer instructions and accepted procedures associated with the selected analytical methods, the respective laboratory's QAPP, and SOPs.

All instruments will be calibrated to known, traceable standards. Typically, certified standards with certificates of analysis are used to prepare calibration standards for analytical instruments. Certified calibration standards are obtained from commercial vendors, and where possible, standards be traceable to the National Institute of Standards and Technology (NIST). Stock solutions for spiking solutions will be made from reagent-grade chemicals or as specified in the SOPs. All new calibration or spiking solutions are analyzed against a previously accepted standard to verify that the concentrations of each parameter are within 15% of the verified stock.

Preparation of a standard curve is accomplished by using calibration standards containing the analyte to be measured into a specific solvent mixture to be introduced into the instrument. The concentrations of the working standards are chosen to cover the working range of the instrument. The calibration curve is prepared by plotting instrument response versus the concentration of the standards. Concentrations of the samples analyzed are read directly from the calibration curve or determined by interpolation.

Calibration procedures, checks, and results will be recorded in field and laboratory logbooks. Testing will not occur until instrument calibration results meet the acceptance criteria.

B.8. Inspection/Acceptance of Supplies and Consumables

Standard operating procedures itemize the apparatus, equipment, materials, and supplies required for various field and laboratory equipment, and for each analytical technique. In general, supplies and consumables are procured directly from the vendor. Supplies must meet the following criteria:

- solvent and reagent grades are based on the intended use. All materials must meet the purity requirements of the method;
- equipment used to generate data must provide appropriate sensitivity;
- a certificate of analysis must be provided and retained for reagents and standards;
- the quality and purity of expendable materials must be documented and adequate to meet the data quality objectives of the client.

All reagents, calibration standards, chemicals, and other supplies are to be inspected upon receipt by qualified staff to ensure suitability for the BWTS tests. No standards, solutions, buffers, or other chemical additives will be used if the expiration date has passed. In the case of field and laboratory equipment and materials, the received item(s) are inspected to ensure the proper part number was received as ordered and to identify any damaged products. If damaged or inappropriate goods are received, they will be returned or disposed of and arrangements will be made to receive replacement materials.

Materials and supplies received are dated so that storage duration can easily be determined. A revolving inventory system (first in, first out) is used to ensure that storage times do not affect the material's integrity. If a manufacturer or SOP requirement indicates a specific expiration period for supplies, those supplies exceeding expiration dates are discarded if not used within the acceptable period.

It is the responsibility of the PC to keep appropriate records, such as logbook entries or checklists, to verify the inspection/acceptance of supplies and consumables, and restock these supplies and consumables when necessary. MERC partner and any other sub-contracted laboratories will follow procedures in their laboratory's QAPP and SOPs for inspection/acceptance of supplies and consumables.

B.9. Non-Direct Measurements

All BWTS test data will be generated through MERC field activities and consequent lab analyses. Several types of data and information may be obtained from other sources for use in support of the tests, such as historical monitoring data from Baltimore Harbor and other test sites, up-to-date equipment manufacturers' operational literature, and National Weather Service data. Non-direct measurement data will be evaluated, documented, and referenced in any document for which they are used in accordance with the EPA document *Guidance for Quality Assurance Project Plans*, EPA/QA G-5, December, 2002.

B.10. Data Management

Data management encompasses and traces the path of the data from their generation to their final use or storage (e.g., from field measurements and sample collection/recording through transfer of data to computers (laptops, data acquisition systems, etc.), laboratory analysis, data validation/verification, QA assessments and reporting of data of known quality to the clients and sponsors. It also includes control mechanisms for detecting and correcting errors.

Various types of data will be acquired and recorded electronically or manually by MERC staff during a BWTS test. Sample collection data (e.g., date, time, and location of collected samples), water quality and chemistry analysis data (e.g., TSS, TOC, and active substance concentration), microbial analysis data (e.g., sample preparation, incubation, and direct counts), phytoplankton analysis data (e.g., number of live and number of dead entities), zooplankton analysis data (e.g., sample concentration; number of dead, total, and live organisms), and whole effluent toxicity test data (e.g., test set up, direct counts, and test take down) are recorded by hand (using indelible ink) on pre-printed data collection forms and/or in bound laboratory notebooks that are uniquely-identified and are specific to the treatment technology being tested. As soon after collection as possible, field notes, data sheets, core logs, and chain-of-custody forms will be scanned to create an electronic record. Biological and chemical data that are recorded by hand are manually entered into either a Microsoft (MS) Access Database or the data are entered into a MS Excel Spreadsheet.

In-tank water quality data (e.g., temperature, pH, dissolved oxygen, salinity, turbidity, and chlorophyll) is automatically recorded in a MS Excel spreadsheet. Facility data (e.g., flow rates and pressure measurements) are electronically recorded every 15 seconds during intake and discharge. These data are exported to MS Excel for subsequent analysis. Results from the laboratory analytical instruments are compiled by laboratory staff in electronic format and submitted to the MERC PC upon obtaining results before the beginning of each test run.

Records received by or generated by any of the MERC staff during the BWTS test will be reviewed by the Data Manager or PC within 2 weeks of receipt or generation, respectively, before the records are used to calculate, evaluate, or report test results. The review will be documented as the dated initials of the reviewer.

All electronic testing records and documents and data files will be stored on a test-specific MERC secured Local Area Network (LAN) that can be accessed only by relevant MERC personnel. The MERC Data

Manager is the single point of control for access to the LAN. The LAN is automatically backed up every 24 hours. The electronic data files are also stored on the MERC internal website, which acts as a secondary data backup/storage mechanism site. Once testing is complete, all testing records and documents are sent to MERC for archival within 2 months of project close-out.

Various personnel are responsible for separate or discrete parts of the data management process:

- MERC field staff are responsible for field measurements/sample collection and recording of data and subsequent shipment of samples to laboratories for analyses. They assemble data files, which includes raw data, calibration information and certificates, QC checks (routine checks), data flags, sampler comments and meta data where available. These files are assembled and forwarded for secondary data review by the MERC Data Manager.
- Laboratories are responsible to comply with the data quality objectives specified in the QAPP and as specified in the laboratory QAP and method specific SOPs. Validated sample laboratory data results are reported to the MERC Data Manager and appropriate senior researcher.
- Secondary reviewers (MERC PC, QA Manager) are responsible for the QC the review, verification and validation of field and laboratory data and reporting validated data to the MERC Director.
- The MERC QA Manager is responsible for performing routine independent reviews of data to ensure the monitoring projects data quality objectives are being met. Findings and recommended corrective actions (as appropriate) are reported directly to project management.
- The MERC Director is responsible for final data certification

C. ASSESMENT AND OVERSIGHT

Internal and external assessments will be conducted for each BWTS test in accordance with Section 10 of the MERC QMP to ensure that:

- elements of this QAPP and specific Test Plans have been correctly implemented;
- the quality of the data generated is adequate and satisfies the DQOs identified in the QAPP;
- corrective actions, when needed, are implemented in a timely manner, and their effectiveness is confirmed.

C.1. Assessment and Response Actions

One of the major objectives of the QAPP is to establish mechanisms necessary to anticipate and resolve potential problems before the quality of performance is compromised. Internal QC measures described in this QAPP, which is peer reviewed by a panel of outside experts, implemented by the technical staff, and monitored by the MERC PC, will yield day-to-day information on data quality. The responsibility for interpreting the results of these checks and resolving any potential problems resides with the MERC PC. Technical staff has the responsibility to identify problems that could affect data quality or the ability to use the data. Any problems that are identified will be reported to the MERC PC, who will work with the MERC QA Manager to resolve any issues. Action will be taken to identify and appropriately address the issue and minimize losses and correct data, where possible. The MERC PC will also relay testing progress and data to the MERC Director on a weekly basis. MERC will be responsible for ensuring that the audits described in the following subsections are conducted as part of each BWTS test.

The standard oversight mechanisms for a BWTS test include: (1) performance evaluation audits (PEAs); (2) technical system audits (TSAs); data quality audits (DQAs). A general test schedule of audits is given in Table 9.

C.1.1. Performance Evaluation Audits

PEAs are described in Section 10.2.3.1 of the MERC QMP. A PEA will be conducted to assess the quality of the variable measurements made in each BWTS test. A PEA involves submitting performance evaluation samples for analysis for each analytical method used in the test. The PE samples contain analytes of interest for the test, preferably in the anticipated concentrations of field samples. The PEA answers questions about whether the measurement system is operating within control limits and whether the data produced meet the analytical QA specifications.

The MERC QA Manager will submit "blind" PE samples to the laboratory as part of a field-sampling event. The QA Manager will evaluate the results of the PE samples as soon as they are received from the laboratory. The critical elements for review of PE results include:

- correct identification and quantitation of the PE sample analytes;
- accurate and complete reporting of the results;
- measurement system operation within established control limits for precision and accuracy.

The concentrations reported for the PE samples will be compared to the known or expected concentrations spiked in the samples. The percent recovery will be calculated and the results compared to the accuracy criteria. The QA Manager will submit the results of the PE samples and the associated calibration and quality control data to the MERC PC and Director. If the accuracy criteria are not met, the laboratory will investigate the cause of the failure and submit a corrective action report to the QA Manager, PC, and MERC Director.

C.1.2. Technical Systems Audits

TSAs are described in Section 10.2.3.2 of the MERC QMP. The MERC QA Manager will perform a TSA at least once during each BWTS test. The purpose of this audit is to ensure that the tests are being performed in accordance with the MERC QMP, this QAPP, and the specific Test Plan.

During this audit, the MERC QA Manager will compare actual test procedures to those specified or referenced in this plan and review data acquisition and handling procedures. The QA Manager will include a review of the testing facility, equipment (calibration, maintenance, and operation) and observation of testing and records (including custody forms). The QA Manager also will check data acquisition procedures and may confer with the vendor and technical staff. The TSA will be guided by a project-specific checklist based on this QAPP.

A TSA report will be prepared as a memo to the MERC PC within 10 business days after completion of the audit; the completed checklist will be attached. The MERC Director will be copied on the memo. The MERC PC will respond to the audit within 10 business days. The MERC QA Manager will verify that all audit Findings and Observations have been addressed and that corrective actions are appropriately implemented. A copy of the complete TSA report with corrective actions will be provided to the MERC Director within 10 business days after receipt of the audit response. The TSA findings will be communicated to technical staff at the time of the audit and documented in a TSA report.

C.1.3. Data Quality Audits

The MERC QA Manager will audit at least 20% of the sample results acquired in the BWTS test and 100% of the calibration and QC data per the QAPP requirements. A checklist based on the QAPP will

guide the audit. An initial ADQ will be conducted on the first batch of test data within 10 business days of when data are available to identify errors early in the data reduction process. The first batch is defined as the testing and variable data generated over the first 2 weeks of testing by the MERC PC. The remaining data will be audited after all data for a technology has been posted on the project SharePoint site and once all statistical analyses are complete. The primary focus of this second audit will be the variable data. Finally, a third ADQ will trace the data from initial acquisition, through reduction and statistical comparisons, to final presentation in the reports and verification statements. It will also confirm reconciliation of the two ADQs.

All formulae applied to the data will be verified, and 20% of the calculations will be checked. Data for the technologies will be reviewed for calculation and transcription errors and data traceability. An audit report will be prepared as a memo to the MERC PC within 10 business days after completion of each data audit; the completed checklist will be attached. The MERC Director will be copied on the memo. The MERC PC will respond to the audit within 10 business days. The MERC QA Manager will verify that all audit Findings and Observations have been addressed and that corrective actions are appropriately implemented. A copy of the complete ADQ report with corrective actions will be provided to the MERC Director within 10 business days after receipt of the audit response.

C.1.4. QA/QC Reporting

Each assessment and audit will be documented in accordance with Section 3.3.4 of the MERC QMP. The results of the PEA, including raw data and calculations, will be reported as stated in Section C1.1. The results of the TSA and ADQ will be submitted to the MERC Director. Assessment reports will include the following:

- identification of Findings and Observations;
- recommendations for resolving problems;
- response to adverse findings or potential problems;
- confirmation that solutions have been implemented and are effective;
- citation of any noteworthy practices that may be of use to others.

C.2. Reports to Management

During the BWTS evaluation, any deviations from this QAPP or the specific Test Plan will be reported immediately to the MERC Director. The MERC QA Manager, during the course of any assessment or audit, will identify to the technical staff performing experimental activities any immediate corrective action that should be taken. A summary of the required assessments and audits, including a listing of responsibilities and reporting timeframes, is included in Table 9.

Corrective action procedures for a specific BWTS test depend on the severity of the nonconformance condition. In cases in which immediate and complete corrective action is implemented by project personnel, the corrective action will be recorded in the appropriate log book. Non-conformance conditions which could have an impact on project data quality must be communicated within 24 hours to the MERC PC and the MERC Director, who is authorized to stop work. These types of issues require a formal corrective action and root cause analysis. The PIs, or MERC QA Manager, can require laboratory activities to be limited or discontinued until the corrective action is complete and the non-conformance issue has been eliminated. Laboratory corrective action procedures are defined in each participating laboratory's QA manual and SOPs. The PI at each organization is responsible for verifying that corrective action is implemented according to internal laboratory policies, this QAPP, and the specific test QA/plan. The individual laboratory PIs are responsible for investigating and implementing test-level corrective actions to address errors or deviations in the laboratory.

Once the TSA or ADQ report has been prepared, the PC will respond to each Finding or Observation following the timeline defined in section C1 and will implement any necessary corrective action. The MERC QA Manager will verify that corrective action has been implemented effectively.

In addition to this QAPP and the specific Test Plan, a final report for each participating vendor and for each technology verified will be prepared and reviewed. The final report is a comprehensive document describing the BWTS test and will include all technologies from a particular vendor. Each draft report will be submitted to the respective vendor for review as well as the expert peer reviewers. Each final report will be signed by the MERC Director. Original signed reports will be provided to the respective vendors and posted on the MERC website (www.maritime-enviro.org).

Table 9. Summary of assessment reports[1].

Assessment	Prepared By	Report Submission Timeframe	Submitted To
TSA	MERC QA Manager	TSA response is due to QA Manager within 10 business days. TSA responses will be verified by the QA Manager and provided to MERC Director within 20 business days.	MERC Director
ADQ 1 (first batch)	MERC QA Manager	ADQ will be completed within 10 business days after receipt of first data set	MERC Director
ADQ 2 (raw data)	MERC QA Manager	ADQ will be completed once all data are received and analyzed	MERC Director
ADQ 3 (synthesized data and final report)	MERC QA Manager	ADQ will be completed within 10 business days after completion of the final report review	MERC Director

[1]Any QA checklists prepared to guide audits will be provided with the audit report.

D. DATA VALIDATION AND USABILITY

All data produced under this QAPP will be evaluated to determine the validity of the on-site monitoring events and the laboratory analyses. Each MERC field and laboratory team member is responsible for ensuring that all records and results they produce or handle are completely and correctly recorded, transcribed, and transmitted. They also are responsible for ensuring that all activities performed) comply with all requirements outlined in this QAPP, the MERC QMP, and relevant SOPs. The MERC is responsible for final verification and validation of all results.

D.1. Data Review, Verification, and Validation Requirements

Data review is the in-house examination to ensure that the data have been recorded, transmitted, and processed correctly. It includes completeness checks to determine if there are any deficiencies such as missing data or lost integrity. Data verification is the process for evaluating the completeness, correctness, and conformance / compliance of the data set against method, procedural and contractual specifications. Data validation is an analyte and sample specific process to determine the quality of a specific data set relative to the end use.

The key data review requirements for the performance evaluation test are the collection of QC samples as outlined in the QAPP, a comparison of raw data sheets and comments against final data to flag any suspect data, and a review of final data to resolve any questions about apparent outliers. The QA audits, as described within this document are designed to assure the quality of these data. In general, the data generated during this test will be reviewed by a MERC technical staff member within 5 days of generation of the data. The reviewer will be familiar with the technical aspects of the BWTS test but will not be the person who generated the data. This process will ensure that the data have been recorded, transmitted, and processed properly. Furthermore, this process will ensure that the monitoring systems data were collected under appropriate testing conditions.

The key data verification requirements for this test are stated in Section B10 of this QAPP. Data verification will be performed by the MERC PC and QA Manager. The MERC Director will be notified by the QA Manager when inconsistencies or non-compliant monitoring and/or laboratory data are discovered. For field activities, it is necessary to determine whether the samples/monitoring data were collected using the sampling/monitoring design specified in section B1, whether the samples/monitoring events were collected according to a specific method or SOP as specified in Section B2, whether the collected samples have been recorded and handled properly as in Section B3, and whether the proper amount of QC samples and procedures were taken to satisfy the QC requirements specified in Section B5. For analytical activities, each sample/monitoring event should be verified to ensure that the procedures used to generate the data (as specified in Section B4) were performed as specified. The proper amount of QC checks (as specified in Section B5) that were prepared and analyzed during the actual analysis provide an indication of the quality of the data. Instrument calibrations (as specified in Section B7) are evaluated to determine whether the correct number of calibration standards were used and the range of the analysis, whether standards were analyzed in an appropriate sequence specific to the methods used, and were performed prior to monitoring events and analysis of samples, blanks, and QC samples in an appropriate time frame.

The MERC PC and QA Manager are responsible for assessing the data against a set of criteria to verify its validity prior to use. The data validation process summarizes the data and QC deficiencies, and determines the impact on the overall data quality. Data validation qualifiers (DQIs and MQOs) are assigned in the data assessment records, flagged on the results tables, and noted in the "Results" section of the final evaluation reports.

D.2. Verification and Validation Methods

Data verification is conducted as part of the data review as described in Section B10 of this QAPP. A visual inspection of handwritten data will be conducted at each level in the field and in the laboratory to ensure that all entries were properly recorded or transcribed and any erroneous entries were properly noted. Records produced electronically or maintained as hard copies are subject to data verification. During field activities, records associated with monitoring events and sample collection such as field data sheets, COC records, shipper's air bills, logbook documentation, or electronic devices to log samples or print sample labels are verified against approved SOPs or procedures. At sample receipt, COC records are verified along with refrigerator and freezer logs to ensure the integrity of the samples. During the sample preparation; certificates of analysis for surrogates and spiking compounds, refrigerator and freezer logs, analytical requests and standard preparation logs are verified. Manufacturer's certificates for calibration and/or internal standards, instrument run or injection logs, standard preparation logs, calculation worksheets, and QC monitoring events/sample results are verified during the analysis of the sample set.

All calculations used to transform the data will be reviewed to ensure the accuracy and the appropriateness of the calculations. Calculations performed manually will be reviewed and repeated using

a handheld calculator or commercial software (e.g., Excel). Calculations performed using standard commercial office software (e.g., Excel) will be reviewed by inspection of the equations used for the calculations and verification of selected calculations by handheld calculator. Calculations performed using specialized commercial software (i.e., for analytical instrumentation) will be reviewed by inspection and, when feasible, verified by handheld calculator, or standard commercial office software.

To ensure that the data generated from any BWTS test meet the goals of the test, a number of data validation procedures will be performed by a MERC team member not immediately responsible for the generation of the data. Sections B and C of this QAPP provide a description of the validation safeguards employed for each BWTS test. Data validation efforts include the completion of QC activities and the performance of a TSA audit as described in Section C. The data from this test will be evaluated relative to the measurements to ensure that the DQOs are met. Data failing to meet these criteria will be flagged in the data set and not used for evaluation of the technologies, unless these deviations are accompanied by descriptions of their potential impacts on the data quality. An ADQ will be conducted by the MERC QA Manager to ensure that data review, verification, and validation procedures were completed and to assure the overall quality of the data.

D.3. Reconciliation with User Requirements

The purpose of a test performed following this QAPP is to evaluate the performance of BWTS. To meet the requirements of the user community, the data obtained in such an evaluation will include thorough documentation of the technology's performance during the test. The data review, verification, and validation procedures described above will assure that test data meet these requirements, are accurately presented in the reports generated from the test, and that data not meeting these requirements are appropriately flagged and discussed in the reports. Additionally, all data generated from any reference method used to evaluate technology results during the test should meet the QA requirements of any applicable standard operating procedures or instrumentation instruction manuals.

This QAPP and the resulting MERC evaluation report(s) will be subjected to review by the participating BWT vendors, MERC staff and Co-PIs, MARAD, and external expert peer reviewers. These reviews will assure that this QAPP and the resulting evaluation report(s) meet the needs of potential users of the BWTS. Performance data for the BWTS, collected under conditions where the QC requirements for the duplicate and PEA samples were met, will be presented in the final evaluation report without any further comment. The final evaluation report(s) will be submitted to the Maryland Port Administration and MARAD in Adobe portable document format (pdf), and subsequently posted on the MERC website.

E. REFERENCES

American National Standards Institute and American Society for Quality Control (ANSI/ASQC), 2004. Quality Systems for Environmental Data Collection and Environmental Technology Programs: Requirements with Guidance for Use. E4-2004. Milwaukee, WI.

Environmental Protection Agency (EPA). 2002. Guidance for Quality Assurance Project Plans, EPA QA/G-5, EPA/240/B-01/003. Office of Environmental Information. Washington, D.C.

Environmental Protection Agency (EPA). 2002. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms. Third edition.. EPA-821-R-02-014. Office of Water, Washington, DC. www.epa.gov/OST/WET/

Environmental Protection Agency (EPA). 2006. Guidance on Systematic Planning Using the Data Quality Objectives Process. EPA QA/G-4. EPA/240/B_06/001. Office of Environmental Information. Washington, D.C

Environmental Protection Agency (EPA). 2010. Generic Protocol for the Verification of Ballast Water Treatment Technology. EPA/600/R-10/146. Washington, D.C.

Hasan, J. A., Huq, A., Tamplin, M. L., Siebeling, R. J. and R.R. Colwell. (1994). A novel kit for rapid detection of *Vibrio cholerae* O1. *J Clin Microbiol*. 32: 249–252.

Hubaux, A. G. and G. Vos. 1970. Decision and Detection Limits for Linear Calibration Curves. Analytical Chemistry, Vol 42, No. 8, pp. 849-855.

International Maritime Organization (IMO). 2004. International Convention for the Control and Management of Ships' Ballast Water and Sediments (BWM). London, www.imo.org.

International Maritime Organization (IMO). 2005. Resolution MEPC.125(53) Guidelines for Approval of Ballast Water Management Systems (G8)

International Maritime Organization (IMO). 2008. Resolution MEPC.125(57) Revised Procedure for Approval of Ballast Water Management Systems that Make Use of Active Substances (G9)

Miller, A.W., M. Frazier, G.E. Smith, E.S. Perry, G.M. Ruiz, and M.N. Tamburri, 2011. Enumerating Sparse Organisms in Ships' Ballast Water: Why Counting to 10 is Difficult? *Environ. Sci. Tech.*, 45:3539-3546.

Steinberg, M.K., E.J. Lemieux, and L.A. Drake, 2011. Determining the viability of marine protists using a combination of vital, fluorescent stains. *Mar Biol* 158:1431-1437.

APPENDIX A

MERC Data Quality Objectives

Table A. MERC Data Quality Objectives for Physical/Chemical Analyses

Data Quality Indicator	Evaluation Process	Performance Measure	MERC DQO
Precision	Analyze at least 10 % of samples in duplicate.	Relative Percent Difference (RPD).	< 20 % average RPD.
Bias	*Experiment Bias:* Analysis of spike-recovery samples; ensure proper calibration/verification and maintenance of equipment/analytical instrumentation; ensure proper sample handling to avoid contamination.	Percent Spike Recovery (SPR)	75%-125% average SPR.
Accuracy	Where applicable, use a certified reference standard to determine differences between the measured and nominal reference standard concentrations.	Percent Difference (%D).	< 20 % average %D.
Representativeness	Ensure pre-treatment/control and post-treatment/treatment samples are handled and analyzed in the same manner.	N/A – Qualitative term.	N/A – Qualitative term.
Comparability	Routine procedures are conducted according to appropriate SOPs to ensure consistency between tests. Ensure correct implementation of SOPs.	N/A – Qualitative term.	N/A – Qualitative term.
Completeness	Calculate percentage of valid samples analyzed out of the total number of samples collected.	Percent Completeness (%C)	Greater than 90 %C.
Sensitivity	Determine the method detection limit and limit of quantification for each analyte and analytical method utilized.	Method Detection Limit (MDL) and Limit of Quantification (LOQ)	Dependent upon the analyte and instrumentation.

Table B. MERC Data Quality Objectives and Criteria for Zooplankton (Organisms >50 μm)

Data Quality Indicator	Evaluation Process	Performance Measure	MERC DQO
Precision	Analyze at least two slides or two counting wheels from every sample collected	Coefficient of variation (%CV).	≤ 20 % CV.
Bias	Experiment Bias: Ensure proper calibration/verification and maintenance of equipment/analytical instrumentation; ensure proper sample handling to avoid contamination.	N/A	N/A
	Operator Bias: Ensure a second, suitably-qualified operator analyzes at least 10 % of treatment discharge samples, and 10 % of control intake and discharge samples.	Percent Similarity (PSC) and Relative Percent Difference (RPD)	≥90% average PSC and ≤20%average RPD
Representativeness	Ensure sample water contains biota representative of harbor water in terms of species composition and richness. Ensure pre-treatment/control and post-treatment/treatment samples are handled and and analyzed in the same manner.	N/A – Qualitative term	N/A – Qualitative term
Comparability	Routine procedures are conducted according to appropriate SOPs to ensure consistency between tests. Ensure correct implementation of SOPs.	N/A – Qualitative term	N/A – Qualitative term
Completeness	Calculate percentage of valid samples analyzed out of the total number of samples collected.	Percent Completeness (%C)	Greater than 90 %C.

Table C. Data Quality Objectives and Criteria for Phytoplankton (i.e., Entities <50 and >10 μm)

Data Quality Indicator	Evaluation Process	Performance Measure	MERC DQO
Precision	Analyze at least two out of five treatment discharge samples and at least one out of five control intake or discharge samples (from each set of five test trials).	Percent Similarity (PSC)	≥ 60% average PSC between paired replicates.
Bias	Experiment Bias: Ensure proper calibration/ verification and maintenance of equipment/analytical instrumentation; ensure proper sample handling to avoid contamination.	N/A	N/A
	Operator Bias: Ensure a second, suitably-qualified operator analyzes at least two treatment discharge samples per set of five test trials, and at least one control intake or discharge sample per set of five test trials.	Percent Similarity (PSC) and Relative Percent Difference (RPD)	≥90% average PSC and ≤20%average RPD
Representativeness	Ensure augmented test organisms are representative of those naturally found in Baltimore Harbor and other test sites. Ensure sample water contains biota representative of test siter water in terms of species composition and richness. Ensure pre-treatment/control and post-treatment/treatment samples are handled and analyzed in the same manner.	N/A – Qualitative term	N/A – Qualitative term
Comparability	Routine procedures are conducted according to appropriate SOPs to ensure consistency between tests. Ensure correct implementation of SOPs.	N/A – Qualitative term	N/A – Qualitative term
Completeness	Calculate percentage of valid samples analyzed out of the total number of samples collected.	Percent Completeness (%C)	Greater than 90 %C.

Table D. Data Quality Objectives and Criteria for Microbial Samples

Data Quality Indicator	Evaluation Process	Performance Measure	MERC DQO
Precision	Analyze at least 10 % of samples in duplicate.	Relative Percent Difference (RPD).	< 20 % average RPD.
Bias	*Experiment Bias:* Ensure proper calibration/verification and maintenance of equipment/analytical instrumentation; ensure proper sample handling to avoid contamination.	N/A	N/A
	Operator Bias: Ensure a second, suitably-qualified operator counts at least 10 % of samples.	Relative Percent Difference (RPD).	< 20 % average RPD.
Representativeness	Ensure sample water contains biota representative of harbor water in terms of species composition and richness. Ensure pre-treatment/control and post-treatment/treatment samples are handled and analyzed in the same manner.	N/A – Qualitative term.	N/A – Qualitative term.
Comparability	Routine procedures are conducted according to appropriate SOPs to ensure consistency between tests. Ensure correct implementation of SOPs.	N/A – Qualitative term.	N/A – Qualitative term.
Completeness	Calculate percentage of valid samples analyzed out of the total number of samples collected.	Percent Completeness (%C)	Greater than 90 %C.
Sensitivity	Limit of detection (LOD) is determined for each analysis type, and reported in the product literature.	Limit of Detection (LOD)	Determined by the manufacturer, dependent on the sample volume analyzed.

Table E. Data Quality Objectives and Criteria for Whole Effluent Toxicity (WET) Testing

Data Quality Indicator	Evaluation Process	Performance Measure	MERC DQO
Bias	*Experiment Bias:* Conduct monthly reference toxicity tests on test organisms or obtain reference toxicity test data from the test organism supplier(s);; ensure proper calibration/verification and maintenance of equipment/analytical instrumentation; ensure proper sample handling to avoid contamination	Determination of the sensitivity of the test organisms relative to historical data using a quality control chart	LC50 value within 2 standard deviations of the historical mean LC50.
	Operator Bias: Ensure a second, suitably-qualified operator counts at least 10 % of samples.	Relative Percent Difference (RPD).	< 10 % average RPD.
Representativeness	Ensure test organisms are representative of those naturally found in Baltimore Harbor, other test sites, and/or recommended by the U.S. EPA. Control groups (reference and dilution control) and treatment groups are handled and analyzed in the same manner.	N/A – Qualitative term.	N/A – Qualitative term.
Comparability	Routine procedures are conducted according to appropriate SOPs to ensure consistency between tests. Ensure correct implementation of SOPs.	N/A – Qualitative term.	N/A – Qualitative term.
Completeness	Calculate percentage of valid samples analyzed out of the total number of samples collected.	Percent Completeness (%C)	Greater than 90 %C.

Table F. MERC Data Quality Objectives for Water Quality Analyses

Data Quality Indicator	Evaluation Process	Performance Measure	MERC DQO
Precision	Analyze at least 10 % of samples in duplicate.	Relative Percent Difference (RPD).	< 20 % average RPD.
Accuracy	Where applicable, use a certified reference standard to determine differences between the measured and nominal reference standard concentrations.	Percent Difference (%D).	< 20 % average %D.
Representativeness	Ensure pre-treatment/control and post-treatment/treatment samples are handled and analyzed in the same manner.	N/A – Qualitative term.	N/A – Qualitative term.
Comparability	Routine procedures are conducted according to appropriate SOPs to ensure consistency between tests. Ensure correct implementation of SOPs.	N/A – Qualitative term.	N/A – Qualitative term.
Completeness	Calculate percentage of valid samples analyzed out of the total number of samples collected.	Percent Completeness (%C)	Greater than 90 %C.

APPENDIX B

Acronyms and Abbreviations

ADQ	Audit of data quality
ANSI	American National Standards Institute
ASQC	American Society for Quality Control
BWTS	Ballast water treatment systems
CBL	University of Maryland Center for Environmental Science's Chesapeake Biological Laboratory
COA	Certificate of analysis
COC	Chain of custody
DQA	Data quality assessment
DQI	Data quality indicator
DQO	Data quality objective
EPA	U.S. Environmental Protection Agency
ETV	Environmental Technology Verification
IMO	International Maritime Organization
LRB	Laboratory record book
MARAD	Maritime Administration
MERC	Maritime Environmental Resource Center
MPA	Maryland Port Authority
MQO	Measurement quality objectives
NASL	Nutrient Analytical Services Laboratory
NIST	National Institute of Standards and Technology
ODU	Old Dominion University
PE A	Performance evaluation audit
QA	Quality assurance
QAAWP	Quality System Annual Report and Work Plan
QAPP	Quality assurance project plan
QC	Quality control
QMP	Quality management plan
QSA	Quality system audit
QSR	Quality system review
SERC	Smithsonian Environmental Research Center
SOP	Standard operating procedure
TSA	Technical systems audit
UMD	University of Maryland College Park
WREC	University of Maryland Wye Research and Education Center

APPENDIX B

Glossary

Acceptance Criteria - Specified limits placed on characteristics of an item, process, or service defined in requirements documents.

Accuracy - A measure of the closeness of an individual measurement or the average of a number of measurements to the true value. Accuracy includes a combination of random error (precision) and systematic error (bias) components that are due to sampling and analytical operations.

Analyte - A specific chemical that can be detected by a given analytical method.

Assessment - The evaluation process used to measure the performance or effectiveness of a system and its elements. As used by MERC, assessment is an all-inclusive term used to denote any of the following: audit, performance evaluation (PE), management systems review (MSR), peer review, inspection, or surveillance.

Audit (quality) - A systematic and independent examination to determine whether quality activities and related results comply with planned arrangements and whether these arrangements are implemented effectively and are suitable to achieve objectives.

Audit of Data Quality (ADQ) - A qualitative and quantitative evaluation of a set of data after it has been collected and 100% verified by project personnel; consisting of tracing at least 10% of the test data from original recording through transferring, calculating, summarizing and reporting to verify that the resulting data are of acceptable quality.

Bias - The systematic or persistent distortion of a measurement process, which causes errors in one direction (i.e., the expected sample measurement is different from the sample's true value).

Blank - A sample subjected to the usual analytical or measurement process to establish a zero baseline or background value. Sometimes used to adjust or correct routine analytical results. A sample that is intended to contain none of the analytes of interest. A blank is used to detect contamination during sample handling preparation and/or analysis. There are many types of blanks, each with a specific purpose including:

> **Equipment Blanks** - Monitor for potential contamination from decontamination procedures of field gear or from other sources of equipment contamination.
> **Field Blank** - A blank used to provide information about contaminants that may be introduced during sample collection, storage, and transport. A clean sample, carried to the sampling site, exposed to sampling conditions, returned to the laboratory, and treated as an environmental sample.
> **Laboratory Blanks** - Samples that are used to identify potential sources of contamination that are generated during the processing and analysis of samples in the laboratory.

Method Blank - A blank prepared to represent the sample matrix as closely as possible and analyzed exactly like the calibration standards, samples, and quality control (QC) samples. Results of method blanks provide an estimate of the within-batch variability of the blank response and an indication of bias introduced by the analytical procedure.

Trip Blank - A clean sample of a matrix that is taken to the sampling site and transported to the laboratory for analysis without having been exposed to sampling procedures.

Chain of Custody - An unbroken trail of accountability that ensures the physical security of samples, data, and records.

Comparability - A measure of the confidence with which one data set or method can be compared to another.

Completeness - A measure of the amount of valid data obtained from a measurement system compared to the amount that was expected to be obtained under correct, normal conditions.

Data Quality Assessment (DQA) - The scientific and statistical evaluation of data to determine if data obtained from environmental operations are of the right type, quality, and quantity to support their intended use. The five steps of the DQA Process include: 1) reviewing the DQOs and sampling design, 2) conducting a preliminary data review, 3) selecting the statistical test, 4) verifying the assumptions of the statistical test, and 5) drawing conclusions from the data

Data Quality Indicators (DQIs) - The quantitative statistics and qualitative descriptors that are used to interpret the degree of acceptability or utility of data to the user. The principal data quality indicators are bias, precision, accuracy (bias is preferred), comparability, completeness, representativeness. DQIs provide a metric against which the performance of a program can be measured during the implementation and/or assessment phases of a verification test.

Data Quality Objectives (DQOs) - The qualitative and quantitative statements derived from the DQO Process that clarify study's technical and quality objectives, define the appropriate type of data, and specify tolerable levels of potential decision errors that will be used as the basis for establishing the quality and quantity of data needed to support decisions.

Data Quality Objectives (DQO) Process - A systematic strategic planning tool that identifies and defines the type, quality, and quantity of data needed to satisfy a specified use. DQOs are the qualitative and quantitative outputs from the DQO Process.

Data Validation - A procedure for assessing whether or not a set of data have met acceptability criteria defined in the data quality objective process.

Detection Limit (DL) - A measure of the capability of an analytical method to distinguish samples that do not contain a specific analyte from samples that contain low concentrations of the analyte; the lowest concentration or amount of the target analyte that can be determined to be different from zero by a single measurement at a stated level of probability. DLs are analyte- and matrix-specific and may be laboratory-dependent. Some of the more commonly used definitions are:

Instrument Detection Limit (IDL) - The lowest concentration or mass an instrument can detect above background instrument noise under ideal conditions. Sample preparation is not considered in the determination of an IDL.

Method Detection Limit - A statistically derived estimate of the lowest concentration or mass detectable under method conditions at the concentration evaluated. A series of standards at an estimated limit of detection is analyzed multiple times (usually 7), a standard deviation of these seven replicate analyses is determined and the standard deviation is multiplied by the Student's t-distribution statistic at 6 degrees of freedom.
Practical Quantitation Limit (PQL) - A measure of the lowest limit of detection under the conditions of a particular method. The PQL is often determined by multiplying the MDL by a factor of between 3 and 10.
Reporting Limit (RL) - For a target analyte, the reporting limit is instrument dependent and based on the lowest point on the calibration curve.

Duplicate Analysis - A measure of precision determined by analyzing samples twice or by analyzing a second sample taken from the same source at the same time and analyzed under identical conditions. There are several different types of duplicate samples that provide information on the precision of specific types of data:

Field Duplicates - Independent samples that are collected as close as possible to the same point in time and space. They are two separate samples taken from the same source, stored in separate containers and analyzed independently. These types of duplicates are useful in characterizing the precision of the sampling process.
Matrix Duplicates - An intra-laboratory split sample that is used to document the precision of a method in a given sample matrix.
Split Samples - Two or more representative portions taken from one sample in the field or in the laboratory and analyzed by different analysts or laboratories. Split samples are quality control (QC) samples that are used to assess analytical variability and comparability.

Duplicate Samples - Two samples taken from and representative of the same population and carried through all steps of the sampling and analytical procedures in an identical manner. Duplicate samples are used to assess variance of the total method, including sampling and analysis.

Existing data - Existing data are data or information that will be used that have not been newly generated by the BWTS test. They may also be known as secondary data or non-direct measurement.

Field (Matrix) Spike - A sample prepared at the sampling point (i.e., in the field) by adding a known mass of the target analyte to a specified amount of the sample. Field matrix spikes are used, for example, to determine the effect of the sample preservation, shipment, storage, and preparation on analyte recovery efficiency (the analytical bias).

Performance Evaluation (PE) - A type of audit in which the quantitative data generated in a measurement system are obtained independently and compared with routinely obtained data to evaluate the proficiency of an analyst or laboratory.

Precision - A measure of mutual agreement among individual measurements of the same property, usually under prescribed similar conditions expressed generally in terms of the standard deviation.

Quality Assurance (QA) - An integrated system of management activities involving planning, implementation, assessment, reporting, and quality improvement to ensure that a process, item, or service is of the type and quality needed and expected.

Quality Assurance Project Plan (QAPP) - A formal document describing in comprehensive detail the necessary quality assurance (QA), quality control (QC), and other technical activities that must be

implemented to ensure that the results of the work performed will satisfy the stated performance criteria. The QAPP components are divided into four classes: 1) Project Management, 2) Measurement/Data Acquisition, 3) Assessment/Oversight, and 4) Data Validation and Usability.

Quality Control (QC) - The overall system of technical activities that measures the attributes and performance of a process, item, or service against defined standards to verify that they meet the stated requirements; operational techniques and activities that are used to fulfill requirements for quality.

Quality Management Plan (QMP) - A formal document that describes the MERC quality system. It is the "blueprint" that defines MERC's QA policies and procedures; the criteria for and areas of QA application; and the different QA-related roles, responsibilities, and authorities of personnel.

Quality System - A structured and documented management system describing the policies, objectives, principles, organizational authority, responsibilities, accountability, and implementation plan of an organization for ensuring quality in its work processes, products (items), and services. The quality system provides the framework for planning, implementing, and assessing work performed by the organization and for carrying out required quality assurance (QA) and quality control (QC).

Quality system audit - An on-site review of the implementation of a verification organization's quality system as documented in the approved QMP. This review is used to verify the existence of, and evaluate the adequacy of, the internal quality system. It is the qualitative assessment of data collection operations and/or organization(s) to evaluate the adequacy of the prevailing quality management structure, policies, practices, and procedures for obtaining the type and quality of data needed.

Raw data - All data and information recorded in support of analytical and process measurements made during planning, testing, and assessing environmental technology including records such as: computer printouts, instrument run charts, standards preparation records, field log records, technology operation logs, and monitoring records.

Record - A statement of data and facts pertaining to a specific event, process, or product, that provides objective evidence that an activity has occurred. All books, papers, maps, photographs, machine readable materials, or other documentary materials, regardless of physical form or characteristics made or received by MERC. Examples include raw and summary data tables, data notebooks, audit reports, and meeting minutes.

Repeatability - The degree of agreement between independent test results produced by the same analyst, using the same test method and equipment on random aliquots of the same sample within a short time period.

Representativeness - A measure of the degree to which data accurately and precisely represent a characteristic of a population, a parameter variation at a sampling point, a process condition, or an environmental condition.

Sensitivity - The capability of a method or instrument to discriminate between measurement responses representing different levels of a variable of interest

Standard Operating Procedure (SOP) - A written document that details the method for an operation, analysis, or action with thoroughly prescribed techniques and steps and that is officially approved as the method for performing certain routine or repetitive tasks.

Technical Systems Audit (TSA) - A thorough, systematic, on-site qualitative audit of sampling and/or measurement systems. The objective of the TSA is to assess and document acceptability of all facilities, equipment, personnel, training, sampling and analytical activities, record keeping, data validation, data management, and reporting aspects of a system. An approved test plan provides the basis for the TSA.
.

Test Plan - The plan developed by MERC for each individual test of a technology. The test plan may include more than one technology. The test/QA plan provides the experimental approach with clearly stated test objectives and associated quality objectives for the related measurements. The test plan may incorporate or reference existing QAPPs.